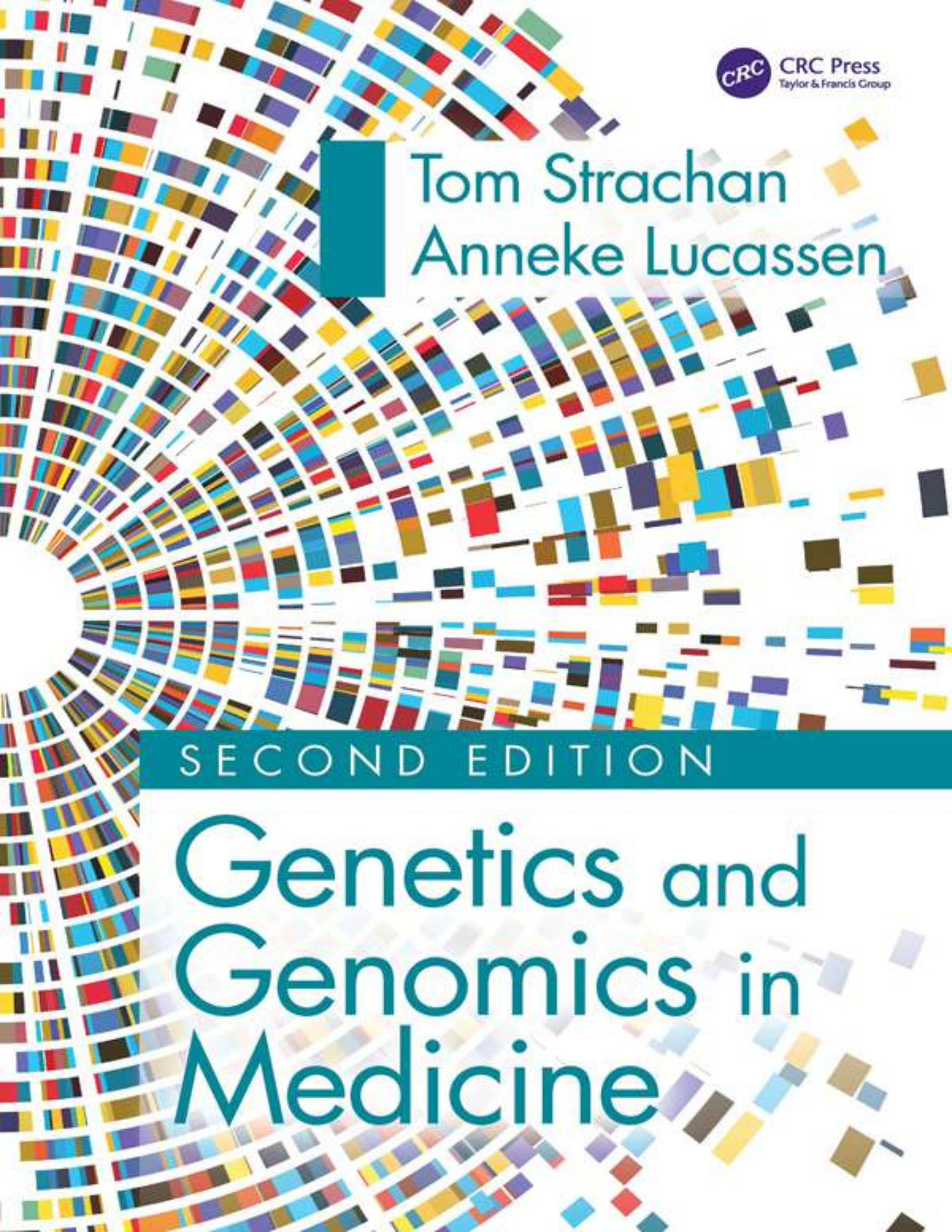
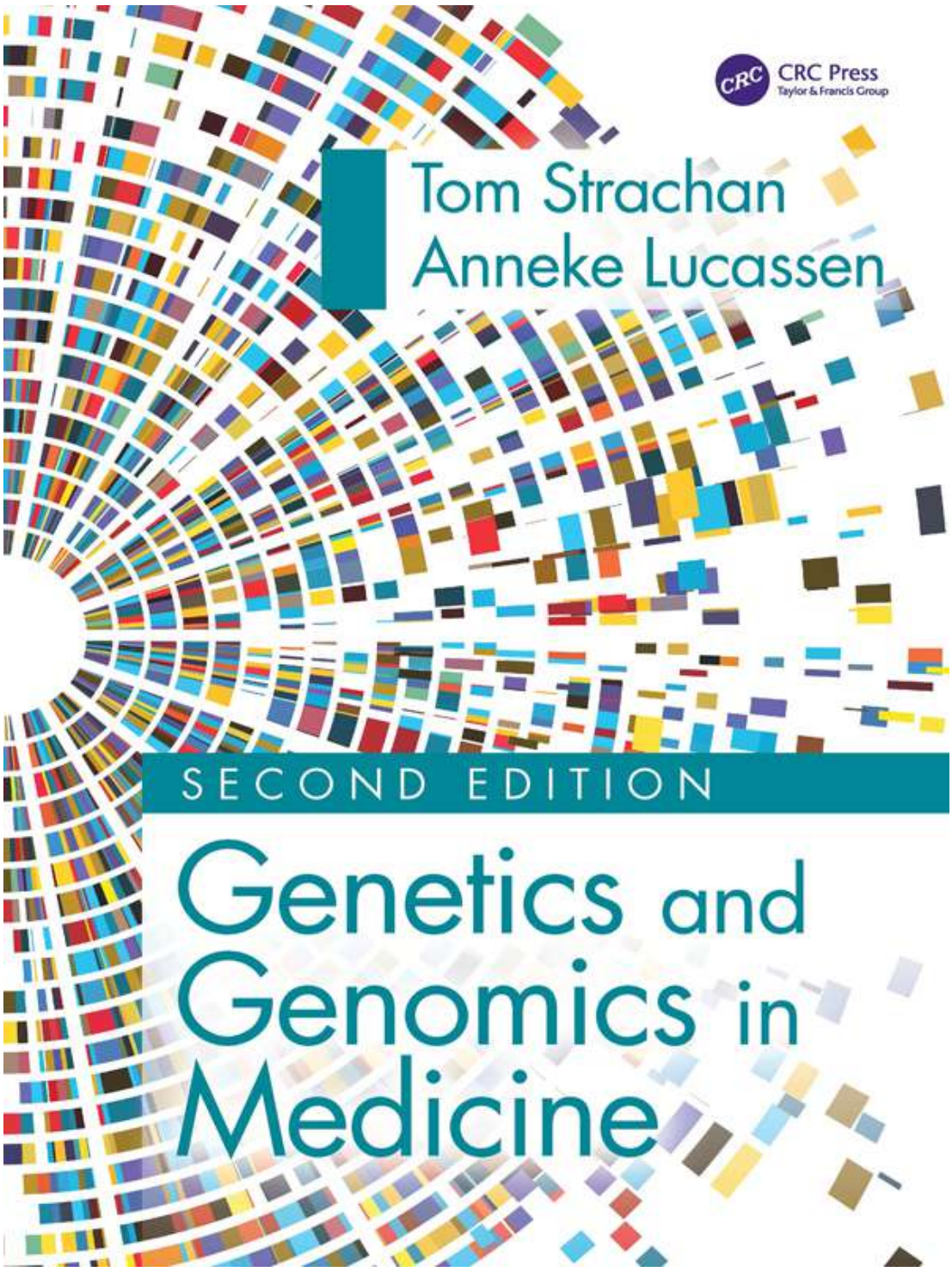

CRC Press
Taylor & Francis Group
Tom Strachan
Anneke Lucassen
SECOND EDITION
Genetics and Genomics in Medicine


CRC Press
Taylor & Francis Group
Tom Strachan
Anneke Lucassen
SECOND EDITION
Genetics and Genomics in Medicine

GENETICS AND GENOMICS IN MEDICINE

SECOND EDITION

GENETICS AND GENOMICS IN MEDICINE

TOM STRACHAN AND ANNEKE LUCASSEN

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Second edition published 2023
by CRC Press
6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

ISBN: 978-0-367-49082-9 (hbk)
ISBN: 978-0-367-49081-2 (pbk)
ISBN: 978-1-003-04440-6 (ebk)

DOI: 10.1201/9781003044406

Typeset in Utopia
by Apex CoVantage, LLC

Access the Support Material at: https://www.routledge.com/9780367490812

Contents

Preface

A rationale for establishing the first edition of *Genetics and Genomics in Medicine* was the suspicion that genomewide analyses might transform medicine. Using Sanger dideoxy sequencing the international Human Genome Project took about 13 years to deliver an almost complete genome sequence in 2003. Subsequent technological developments—first, genome-wide microarray technologies and then massively parallel DNA sequencing—have certainly transformed genome analysis, permitting genome data in hours, not years.

The preface to the first edition of this book also included this question: might we soon live in societies where genome sequencing of citizens becomes the norm? Well, that day seems much closer now as millions of people have their genome sequenced, and debate has begun on whether population neonatal genome sequencing should be considered. The genome sequencing revolution found early major applications in medical genetics, then hematology and oncology, but is now being increasingly applied across multiple other medical disciplines. Various national genomic medicine initiatives have recently been established and, in 2020, NHS England became the first national health service to offer whole genome sequencing to patients as part of routine care.

In this book we try to summarize pertinent knowledge, and to structure it in the form of principles, rather than seek to compartmentalize information into chapters on topics such as epigenetics, evolutionary genetics, immunogenetics, pharmacogenetics, and so on. To help readers find broad topics that might be dealt with in two or more chapters, we provide a road

map on the inside front cover that charts how some broad themes are distributed between different chapters.

We start with three introductory chapters that provide basic background details. Chapters 1 and 2 cover the fundamentals of DNA, chromosomes, the cell cycle, human genome organization and gene expression. Chapter 3 introduces the basics of three core molecular genetic approaches used to manipulate DNA: DNA amplification (by DNA cloning or PCR), nucleic acid hybridization, and DNA sequencing, but we delay bringing in applications of these fundamental methods until later chapters, setting them against appropriate contexts that directly explain their relevance.

The next three chapters provide some background principles at a higher level. In Chapter 4, we take a broad look at general principles of genetic variation, including DNA repair mechanisms and some detail on functional variation (but we consider how genetic variation contributes to disease in later chapters, notably chapters 7, 8 and 10). Chapter 5 takes a look at how genes are transmitted in families and at allele frequencies in populations. Chapter 6 moves from the basic principles of gene expression covered in chapter 2 to explaining how genes are regulated by a wide range of protein and noncoding RNA regulators, and the central role of regulatory sequences in both DNA and RNA. In this chapter, too, we outline the principles of chromatin modification and epigenetic regulation and explain how aberrant chromatin structure underlies many single gene disorders.

The remainder of the book is largely devoted to clinical applications. We explain in chapter 7 how chromosome abnormalities arise and their consequences, and how mutations and large-scale DNA changes can directly cause disease. In chapter 8, we look at how genes underlying single gene disorders are identified, and also how genetic variants conferring susceptibility to complex diseases are identified. Then we consider the ways in which genetic variants, epigenetic dysregulation and environmental factors all make important contributions to complex diseases. Chapter 9 briefly covers the wide the range of approaches for treating genetic disorders, before examining in detail how genetic approaches are used directly and indirectly in treating disease. In this chapter, too, we examine

how genetic variation affects how we respond to drug treatment. Chapter 10 deals with cancer genetics and genomics and explains how cancers arise from a combination of abnormal genetic variants and epigenetic dysregulation. Finally, Chapter 11 takes a broad look at diagnostic applications (and the exciting applications offered by new genome-wide technologies), plus ethical considerations in diagnosis and in some novel therapies.

Important recent advances have been made in applying genetic and genomic technologies to understanding pathogenesis, and in developing novel genetic testing methods, (including noninvasive ones), and novel treatments. There has been significant improvement, too, in pharmacogenomic approaches and in prenatal and preconception options to avoid serious genetic disease. Now we are no longer bound by the old approach of starting with a phenotype and then searching for a confirmatory genotype but can invert the process to predict phenotypes over a lifetime from a genotype. But challenges remain. Predicting phenotypes over a lifetime from a genotype, for example, is rarely clear-cut; the more we test without medical indications, the less likely we will predict diseases accurately. And, while acquiring genetic and genomic data is no longer the major rate-limiting step it was, data interpretation has become a huge challenge given the inherent complexities of interpreting the 4–5 million variants in a person's genome and their implications for [ill] health.

Mainstreaming of genomic medicine—placing it at the center of healthcare — may be appealing, but its utility can be expected to be limited in the first instance to rare diseases and some easily studied cancers. Complex genetic disease is another matter. Genomewide association studies have undoubtedly been successful, especially in improving our understanding of the molecular pathways in a wide range of complex genetic diseases, but they have their limitations. Increasingly, attention has been devoted to finding rare variants by genomewide sequencing (with considerable recent success in some diseases, such as schizophrenia), and in investigating copy number variants. To properly appreciate the complexity of common genetic disease will require more information, too, from other

approaches, investigating modifier genes, environmental factors and so on, and reliance on phenotyping data from large population biobanks will be important.

The familial nature of much genetic information also poses challenges to many modern healthcare services for which there are no clear off-the-shelf solutions. Confidentiality in medicine remains important, yet shared familial inheritances may need disclosing at times, just as we attempted to trace contacts exposed to COVID-19. Sustainability aspects of long-term mass data storage are yet to be examined in any depth, and the lack of population diversity in most of the world's genomic repositories, and thus our understanding of genomic variation, needs urgent attention.

We have tried to convey the excitement of fast-moving research in genetics and genomics and their clinical applications, while explaining how the progress has been achieved. By weaving the ethical, legal and social aspects inherent in these developments throughout the text we hope to provide the reader with a realistic lens through which to view the promising developments in genetics and genomics. There is a long way to go, notably in understanding complex disease and in developing effective treatments for many disorders. But some impressive recent therapeutic advances, and new technological developments such as the prime editing and base editing refinements to CRISPR-Cas genome editing, have engendered an undeniable sense of excitement and optimism. How far will we move from the commonplace one-size-fits-all approach to disease treatment toward an era of personalized or precision medicine? At the very least, we might expect an era of stratified medicine where, according to the genetic variants exhibited by patients with a specific disease, different medical actions are taken.

We would like to thank the staff at CRC Press and Naughton Project Management Ltd: Jo Koster, Jordan Wearing and Nora Naughton, who have undertaken the job of converting our drafts into the finished product. We are also grateful to our family members: Meryl, Alex, James, Tim, Emily and Isobel for their steadfast support.

LITERATURE ACCESS

We live in a digital age and, accordingly, we have sought to provide electronic access to information. To help readers find references cited under Further Reading we provide the relevant PubMed identification (PMID) numbers for the individual articles—see also the PMID glossary item. We would like to take this opportunity to thank the US National Center for Biotechnology Information (NCBI) for their invaluable PubMed database that is freely available at: http://www.ncbi.nlm.nih.gov/pubmed/. Readers who are interested in new research articles that have emerged since publication of this book, or who might want to study certain areas in depth, may wish to take advantage of literature citation databases such as the freely available Google Scholar database (scholar.google.com).

For background information on single gene disorders, we often provide reference numbers to access OMIM, the Online Mendelian Inheritance in Man database (http://www.omim.org). For the more well-studied of these disorders, individual chapters in the University of Washington's GeneReviews series are highly recommended. They are electronically available at the NCBI's Bookshelf within its PubMed database. For convenience, we have given the PubMed Identifier (PMID) for individual articles that we refer to from the GeneReviews series. Note that all GeneReviews articles can be accessed through PubMed at PMID 20301295, where there is an alphabetic listing of all disorders covered by GeneReviews.

Tom Strachan and Anneke Lucassen

Acknowledgements

In writing this book, we have benefited greatly from the advice of many geneticists, biologists and clinicians. We are also grateful to various colleagues who contributed clinical profiles and/or laboratory data for case studies, or who advised on the contents of chapters and/or commented on some aspects of the text, notably the following: Chiara Bettolo, David Bourn, Gareth Breese, Heather Cordell, Jordi Diaz-Manera, Shaun Haigh, Rachel Horton, Majlinda Lako, Richard Martin, Ciaron McAnulty, Robert McFarland, Sabine Specht, Miranda Splitt, and Volker Straub.

1

Fundamentals of DNA, chromosomes, and cells

DOI: 10.1201/9781003044406-1

CONTENTS

Three structures are the essence of life: cells, chromosomes, and nucleic acids. Cells receive basic sets of instructions from DNA molecules that must also be transmitted to successive generations. And DNA molecules work in the context of larger structures: chromosomes.

Many organisms consist of single cells that can multiply quickly. They are genetically relatively stable, but through changes in their DNA they can adapt rapidly to changes in environmental conditions. Others, including ourselves, animals, plants, and some types of fungi, are multicellular.

Multicellularity offers specialization and complexity: individual cells can be assigned different functions, becoming muscle cells, neurons, or lymphocytes, for example. All the different cells in an individual arise originally from a single cell, and so all nucleated cells carry the same DNA sequences. During development, however, the DNA structure within chromosomes is changed to allow specific changes in gene expression that determine a cell's identity, whether it be a muscle cell or a neuron, for example.

Growth during development and tissue maintenance requires cell division. When a cell divides to produce daughter cells, our chromosomes and the underlying DNA sequences must undergo coordinated duplication and then be carefully segregated to the daughter cells.

Some of our cells can carry our DNA to the next generation. When that happens, chromosomes swap segments and DNA molecules undergo significant changes that make us different from our parents and from other individuals.

1.1 THE STRUCTURE AND FUNCTION OF NUCLEIC ACIDS

General concepts: the genetic material, genomes, and genes

Nucleic acids provide the *genetic material* of cells and viruses. They carry the instructions that enable cells to function in the way that they do and to divide, allowing the growth and reproduction of living organisms. Nucleic acids also control how viruses function and replicate. As we describe later, viruses can be highly efficient at inserting genes into human cells, and modified viruses are widely used in gene therapy.

Nucleic acids are susceptible to small changes in their structure (**mutations**). Occasionally, that can change the instructions that a nucleic acid gives out. The resulting genetic variation, plus mechanisms for shuffling the genetic material from one generation to the next, explains why individual organisms of the same species are nevertheless different from

each other. And genetic variation is the substrate that evolutionary forces work on to produce different species. (But note that the different types of cell in a single multicellular organism cannot be explained by genetic variation—the cells each contain the same DNA and the differences in cell types must arise instead by **epigenetic** mechanisms.)

In all cells the genetic material consists of double-stranded DNA in the form of a double helix. (Viruses are different. Depending on the type of virus, the genetic material may be double-stranded DNA, single-stranded DNA, double-stranded RNA, or single-stranded RNA.) As we describe below, DNA and RNA are highly related nucleic acids. RNA is functionally more versatile than DNA (it is capable of self-replication and individual RNA sequences can also serve as templates to make a protein, or act as regulators of gene expression). RNA is widely believed to have developed at a very early stage in evolution. Subsequently, DNA evolved; being chemically much more stable than RNA, it was more suited to being the store of genetic information in cells.

Genome is the collective term for all the *different* DNA molecules within a cell or organism. In prokaryotes—simple unicellular organisms, such as bacteria, that lack organelles—the genome usually consists of just one type of circular double-stranded DNA molecule that can be quite large and has a small amount of protein attached to it. A very large DNA-protein complex such as this is traditionally described as a **chromosome**.

Eukaryotic cells are more complex and more compartmentalized (containing multiple organelles that serve different functions), and they have multiple different DNA molecules. As we will see below, for example, the cells of a man have 25 different DNA molecules but a woman's cells have a genome made up of 24 types of DNA molecule.

In our cells—and in those of all animals and fungi—the genome is partitioned between the nucleus and the mitochondria. Most of the DNA is found in the nucleus, existing as extremely long linear DNA molecules complexed with a variety of different proteins and some types of RNA to form highly organized chromosomes. However, in mitochondria there is just one type of small circular DNA molecule that is largely devoid of

protein. (In plant cells, chloroplasts also have their own type of small circular DNA molecule.)

Genes are the DNA segments that carry the genetic information to make proteins or functional noncoding RNA molecules within cells. The great bulk of the genes in a eukaryotic cell are found in the chromosomes of the nucleus; just a few genes are found in the small mitochondrial or chloroplast DNA molecules.

The underlying chemistry of nucleic acids

Each nucleic acid strand is a polymer, a long chain containing many sequential copies of a simple repeating unit, a **nucleotide**. Each nucleotide in turn consists of a sugar molecule, to which is attached a nitrogenous base and a phosphate group. In DNA the sugar is deoxyribose, which has five carbon atoms that are labeled 1¢ (one *prime)* to 5¢. It is very closely related to ribose, the sugar molecule found in RNA—the only difference is that a hydroxyl (-OH) group at carbon 2¢ of ribose is replaced by a hydrogen atom in deoxyribose (Figure 1.1).

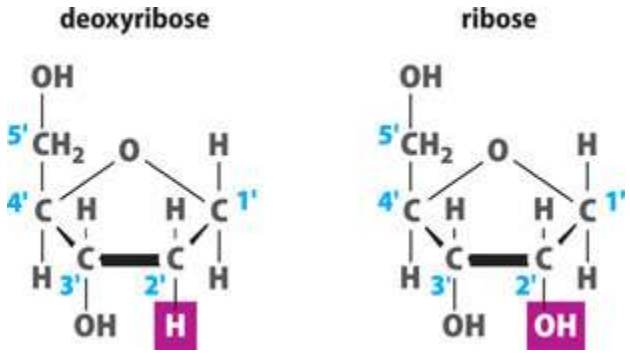


Figure 1.1 Structure of deoxyribose (left) and ribose (right). The five carbon atoms are numbered 1¢ (one *prime)* to 5¢ (five *prime)*. The magenta shading is meant to signify the only structural difference between deoxyribose (the sugar found in DNA) and ribose (the sugar found in RNA): ribose has a hydroxyl (-OH) group in place of the highlighted hydrogen atom attached to carbon 2' of deoyribose. The more precise name for deoxyribose is therefore 2¢-deoxyribose.

Individual nucleotides are joined to their neighbors by a negatively charged phosphate group that links the sugar components of the neighboring nucleotides. As a result, nucleic acids are polyanions, and have a *sugar-phosphate backbone* with bases bonded to the sugars. As explained in **Box 1.1**, the sugar-phosphate backbone of each nucleic acid strand is asymmetric and the ends of each strand are asymmetric, giving direction to each strand.

BOX 1.1 5' AND 3' ENDS, AND STRAND ASYMMETRY OF NUCLEIC ACIDS

In a nucleic acid strand each phosphate group links carbon atom 3¢ from the sugar on one nucleotide to a carbon 5¢ on the sugar of a neighboring nucleotide. Internal nucleotides will therefore be linked through both carbon 5¢ and carbon 3¢ of the sugar to the neighboring nucleotides on opposing sides. However, the nucleotides at the extreme ends of a DNA or RNA strand will have different functional groups. At one end, the **5¢ end**, the nucleotide has a terminal sugar with a carbon 5¢ that is not linked to another nucleotide and is capped by a phosphate group; at the other end, the **3¢ end**, the terminal nucleotide has a sugar with a carbon 3¢ that is capped by a hydroxyl group (**Figure 1**).

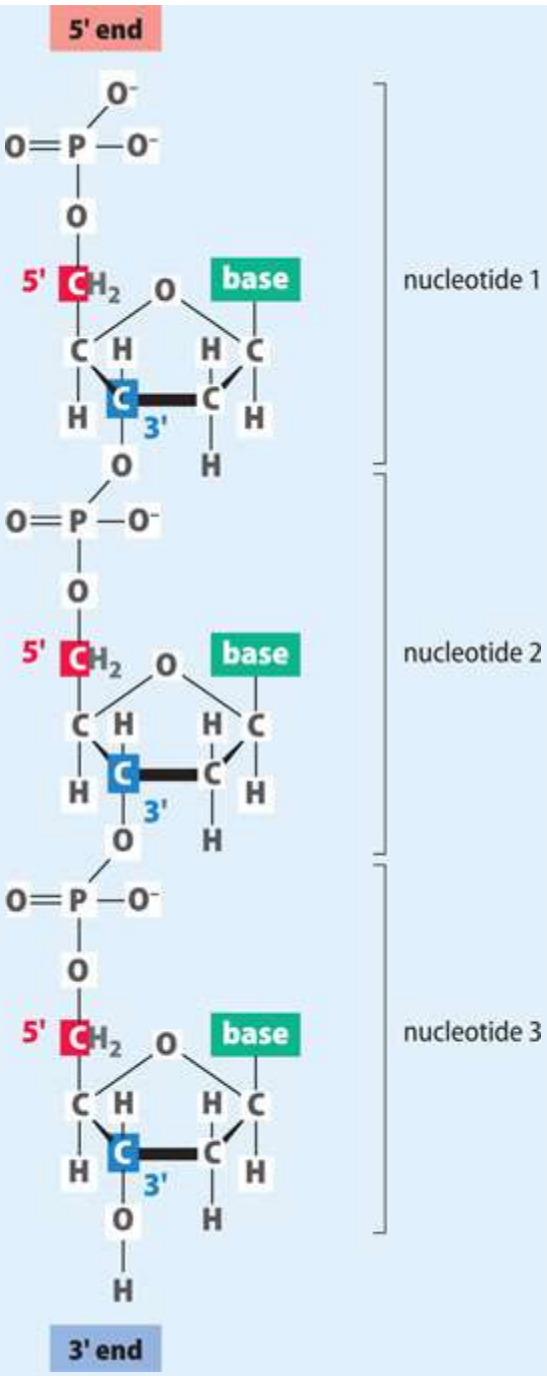


Figure 1 Repeating structure and asymmetric 5′ and 3′ ends in nucleic acids.

The resulting asymmetry between the two ends of a nucleic acid give it a direction. That is important in packing a nucleic acid because when two single nucleic acid strands pair up to make a stable duplex, they must be *anti-parallel*: the 5′ ® 3′ direction of one strand must be opposite to that

of its partner strand. And direction is important for synthesis of a nucleic acid: a growing nucleic acid strand always extends in a 5′ ® 3′ direction.

Unlike the sugar molecules, the nitrogenous bases come in four different types, and it is the sequence of different bases that identifies the nucleic acid and its function. Two of the bases have a single ring based on carbon and nitrogen atoms (a **pyrimidine**) and two have a double ring structure (a **purine**). In DNA the two purines are adenine (A) and guanine (G), and the two pyrimidines are cytosine (C) and thymine (T). The bases of RNA are very similar; the only difference is that in place of thymine there is a very closely related base, uracil (U) (**Figure 1.2**).

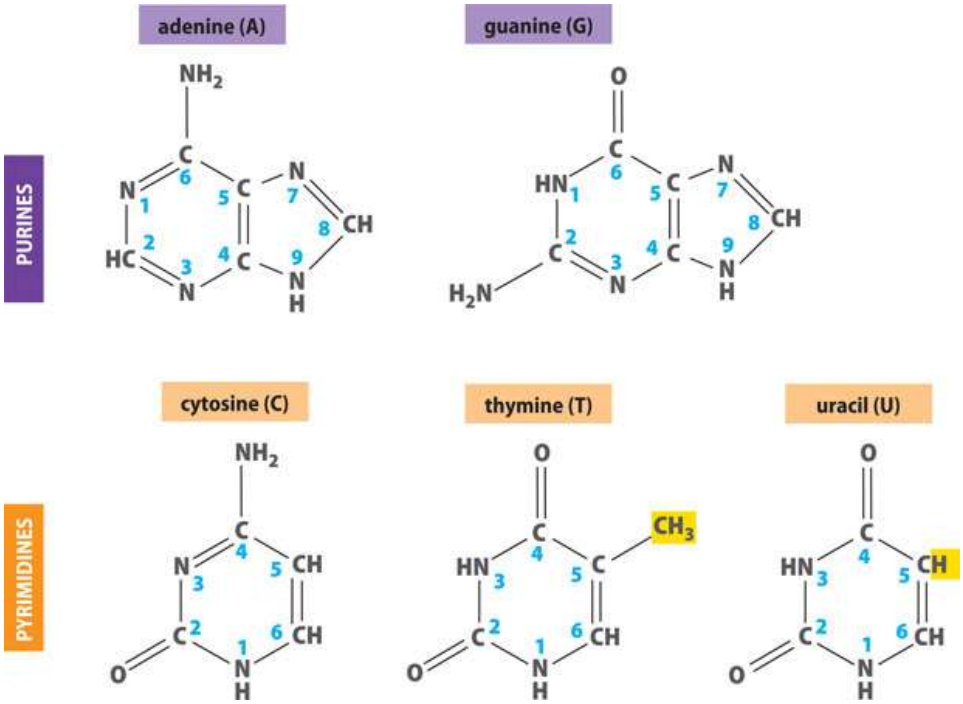


Figure 1.2 Structure of the bases found in nucleic acids. Adenine and guanine are purines with two interlocking rings based on nitrogen and carbon atoms (numbered 1 to 9 as shown). Cytosine and thymine are pyrimidines with a single ring. Adenine, cytosine, and guanine are found in both DNA and RNA, but the fourth base is thymine in DNA and uracil in RNA (they are closely related bases—carbon atom 5 in thymine has an attached methyl group, but in uracil the methyl group is replaced by a hydrogen atom).

Base pairing and the double helix

Cellular DNA exists in a double-stranded (or *duplex*) form, in which the two very long single DNA strands are wrapped round each other. In the resulting double helix each base on one DNA strand is noncovalently linked (by hydrogen bonding) to an opposing base on the opposite DNA strand, forming a **base pair**. However, the two DNA strands fit together correctly only if opposite every A on one strand is a T on the other strand, and opposite every G is a C. (Only two types of base pairs are normally tolerated in double-stranded DNA: A–T and G–C base pairs.) G–C base pairs, which are held together by three hydrogen bonds, are stronger than A–T base pairs, which are held together by two base pairs; see **Figure 1.3**.

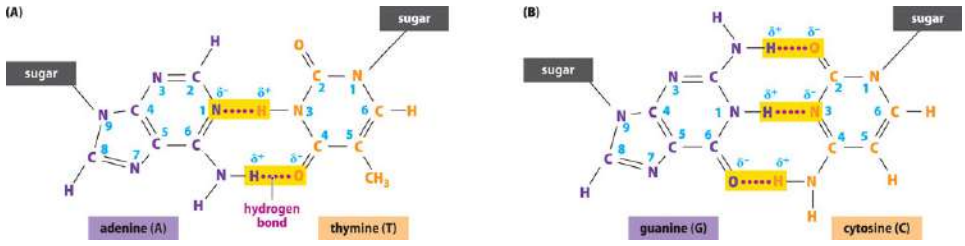


Figure 1.3 Structure of base pairs. In the A–T base pair shown in (A), the adenine is connected to the thymine by two hydrogen bonds. In the G–C base pair shown in (B), three hydrogen bonds link the guanine to the cytosine; a G–C base pair is therefore stronger than an A–T base pair. δ^+ and δ^- indicate fractional positive charges and fractional negative charges.

There is one additional restriction on how two single-stranded nucleic acids form a double-stranded nucleic acid. In addition to a sufficient degree of base pairing, for a duplex to form, the two single strands must be anti-parallel; that is, the 5¢ ® 3¢ direction of one strand is the opposite of the 5¢ ® 3¢ direction of the other strand.

Two single nucleic acid strands that can form a double helix with perfect base matching (according to the base pairing rules given above) are said to have **complementary sequences**. As a result of base pairing rules, the sequence of one DNA strand in a double helix can immediately be used to predict the base sequence of the complementary strand (**Box 1.2**). Note that

base pairing can also occur in RNA; when an RNA strand participates in base pairing, the base pairing rules are more relaxed (see Box 1.2).

DNA replication and DNA polymerases

Base pairing rules also explain the mechanism of DNA replication. In preparation for new DNA synthesis before cell division, each DNA double helix must be unwound using a helicase. During the unwinding process the two individual single DNA strands become available as templates for making complementary DNA strands that are synthesized in the 5¢ ® 3¢ direction (**Figure 1.4**).

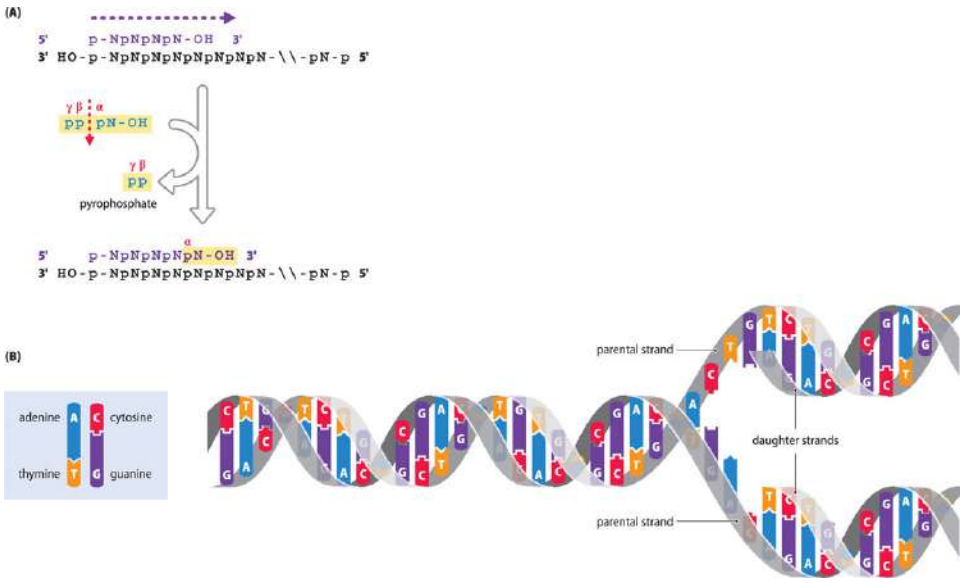


Figure 1.4 DNA synthesis and replication. (A) DNA synthesis. Using a pre-existing DNA strand (black) as a template, a new DNA strand (purple) is synthesized in a 5' ® 3' direction (dashed arrow) using a DNA polymerase to insert successive dNMPs obtained by cleaving the two external phosphates (a and b), to give pyrophosphate residue (which is discarded). (B) DNA replication. The parental DNA duplex consists of two complementary DNA strands that unwind to serve as templates for the synthesis of new complementary DNA strands (daughter strands). Each completed daughter DNA duplex contains one of the two parental DNA strands plus one newly synthesized DNA strand and is structurally identical to the original parental DNA duplex.

BOX 1.2 BASE PAIRING PREVALENCE, SEQUENCE COMPLEMENTARITY, AND SEQUENCE NOTATION FOR NUCLEIC ACIDS

THE PREVALENCE OF BASE PAIRING

The DNA of cells—and of viruses that have a double-stranded DNA genome—occurs naturally as double helices in which base pairing is restricted to A–T and C–G base pairs.

Double-stranded RNA also occurs naturally in the genomes of some kinds of RNA viruses. Although cellular RNA is often single-stranded, it can also participate in base pairing in different ways. Many single-stranded RNAs have sequences that allow intramolecular base pairing—the RNA bends back upon itself to form local double-stranded regions for structural stability and/or for functional reasons. Different RNA molecules can also transiently base pair with each other over short to moderately long regions, allowing functionally important interactions (such as base pairing between messenger RNA and transfer RNA during translation, for example; see Section 2.1). G–U base pairs are allowed in RNA–RNA base pairing, in addition to the standard A–U and C–G base pairs.

RNA–DNA hybrids also form transiently in different circumstances. They occur when a DNA strand is transcribed to give an RNA copy, for example, and when an RNA is reverse transcribed to give a DNA copy.

SEQUENCE COMPLEMENTARITY

Double-helical DNA within cells shows perfect base matching over extremely long distances, and the two

DNA strands within a double helix are said to exhibit **base complementarity** and to have *complementary sequences*. Because of the strict base pairing rules, knowing the base sequence of just one DNA strand is sufficient to immediately predict the sequence of the complementary strand, as illustrated below.

SEQUENCE NOTATION

Because the base sequence of a nucleic acid governs its biological properties it is customary to define a nucleic acid by its base sequence, which is always written in the 5¢ ® 3¢ direction. While a single-stranded oligonucleotide sequence might be written accurately as 5' p-C-p-G-p-A-p-C-p-C-p-A-p-T-OH 3¢, where p = phosphate, it is simpler to write it just as CGACCAT.

For a double-stranded DNA the sequence of just one of the two strands is needed (the sequence of the complementary strand can immediately be predicted by the base pairing rules given above). If a given DNA strand has the sequence CGACCAT, the sequence of the complementary strand can easily be predicted to be ATGGTCG (in the 5¢ ® 3¢ direction as shown below, where A–T base pairs are shown in green and C–G base pairs in blue).

```
5' CGACCAT 3'
   |||||||
3' GCTGGTA 5'
```

DNA replication therefore uses one double helix to make two double helices, each containing one strand from the parental double helix and one newly synthesized strand (semi-conservative DNA replication). Because DNA synthesis occurs only in the 5¢ ® 3¢ direction, one new strand (the *leading strand*) can be synthesized continuously; the other strand (the *lagging strand*) needs to be synthesized in pieces, known as Okazaki fragments (**Figure 1.5**).

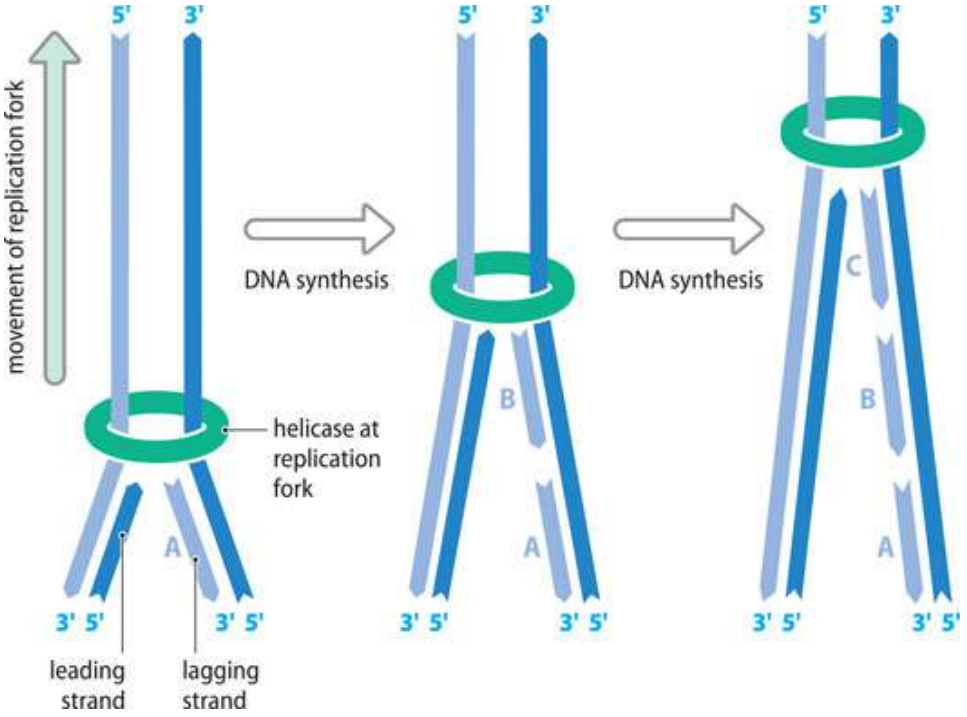


Figure 1.5 Semi-discontinuous DNA replication. The enzyme DNA helicase opens up a **replication fork**, where synthesis of new daughter DNA strands can begin. The overall direction of movement of the replication fork matches that of the continuous 5¢ ® 3¢ synthesis of one daughter DNA strand, the *leading strand.* Replication is semi-discontinuous because the *lagging strand,* which is synthesized in the opposite direction, is built up in pieces (Okazaki fragments, shown here as fragments A, B, and C) that will later be stitched together by a DNA ligase.

Mammalian cells have very many kinds of DNA-dependent DNA polymer-ases that serve a variety of roles, including DNA replication initiation, synthesis of the leading and lagging strands, and also, as described in Section 4.2, multiple roles in DNA repair. Our cells also contain specialized DNA polymerases that use RNA as a template to synthesize a complementary DNA; see **Table1.1**.

TABLE 1.1 CLASSICAL DNA-DEPENDENT AND RNA-DEPENDENT DNA POLYMERASES OF MAMMALIAN CELLS

DNA polymerases	Roles

DNA polymerases	Roles
Classical DNA-dependent DNA polymerases α (alpha) δ (delta) and ε(epsilon) β (beta) γ (gamma)	**Standard DNA replication and/or DNA repair** initiates DNA synthesis (at replication origins, and also when priming the synthesis of Okazaki fragments on the lagging strand) major nuclear DNA polymerases and multiple roles in DNA repair base excision repair (repair of deleted bases and simply modified bases) dedicated to mitochondrial DNA synthesis and mitochondrial DNA repair
RNA-dependent DNA polymerases Retrosposon reverse transcriptase TERT (telomerase reverse transcriptase)	**Genome evolution and telomere function** occasionally converts mRNA and other RNA into complementary DNA, which can integrate elsewhere into the genome; can occasionally give rise to new genes and newexons,and soon. replicates DNA at ends of linear chromosomes, using an RNA template

Note: The classical DNA-dependent DNA polymerases are high-fidelity polymerases—they insert the correct base with high accuracy; however, our cells also have many non-classical DNA-dependent DNA polymerases that exhibit comparatively low fidelity of DNA replication. We will consider the non-classical DNA polymerases in Chapter 4, because of their roles in certain types of DNA repair and in maximizing the variability of immunoglobulins and T-cell receptors.

Genes, transcription, and the central dogma of molecular biology

As a repository of genetic information, DNA must be stably *transmitted* from mother cell to daughter cells, and from individuals to their progeny; DNA replication provides the necessary mechanism. But within the context of individual cells, the genetic information must also be *interpreted* to dictate how cells work. **Genes** are discrete segments of the DNA whose sequences are selected for this purpose, and gene expression is the mechanism whereby genes are used to direct the synthesis of two kinds of product: RNA and proteins.

The first step of gene expression is to use one of the two DNA strands as a template for synthesizing an RNA copy whose sequence is complementary to the selected template DNA strand. This process is called *transcription*, and the initial RNA copy is known as the primary transcript (**Figure 1.6**). Subsequently, the primary transcript undergoes different processing steps, eventually giving a mature RNA that belongs to one of two broad RNA classes:

- *Coding RNA*. RNAs in this class contain sequences that direct the synthesis of polypeptides (the major component of proteins) in a process called translation. This type of RNA has traditionally been called a messenger RNA (mRNA) because it carries genetic instructions to be decoded by the protein synthesis machinery.
- *Noncoding RNA*. All other mature functional RNAs fall into this class, and here the RNAs, not proteins, are the functional endpoint of gene expression. Noncoding RNAs have a variety of roles in cells, as described in later chapters.

In all forms of life, genetic information is interpreted in what initially seemed to be one direction only: DNA ® RNA ® protein, a principle that became known as the central dogma of molecular biology. However, certain DNA polymerases, known as reverse transcriptases, found initially in certain types of viruses, can reverse the flow of genetic information by making a DNA copy of an RNA molecule. The cells of complex organisms also have their reverse transcriptases, as described below. In addition, RNA

can sometimes also be used as a template to make a complementary RNA copy. So, although genetic information in cells mostly flows from DNA to RNA to protein, the central dogma is no longer strictly valid.

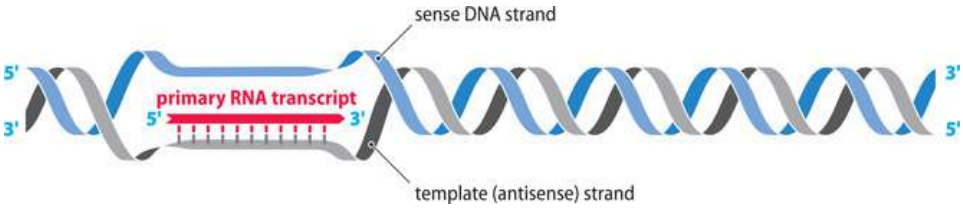


Figure 1.6 Transcription. Transcription results in the synthesis of an RNA transcript in the 5¢ ® 3¢ direction. The nucleotide sequence of the primary RNA transcript is complementary to that of the *template strand* and so is identical to that of the *sense strand,* except that U replaces T. Note: for simplification, the diagram does not show the coiling of the RNA transcript around the template DNA strand to form a double helix.

We will explore gene expression (including protein synthesis) in greater detail in Chapter 2. And in Chapter 6 we will focus on both genetic and epigenetic regulation of gene expression.

1.2 THE STRUCTURE AND FUNCTION OF CHROMOSOMES

In this section we consider general aspects of the structure and function of our chromosomes that are largely shared by the chromosomes of other complex multicellular organisms. We will touch on human chromosomes when we consider aspects of the human genome in Chapter 2, when we first introduce the banded pattern of human chromosomes. In Chapter 7 we consider how disease-causing chromosome abnormalities arise. We describe the methodology and the terminology of human chromosome banding in Box 7.2, and diagnostic chromosome analyses in Chapter 11.

Why we need highly structured chromosomes, and how they are organized

Before replication, each chromosome in the cells of complex multicellular organisms normally contains a single, immensely long DNA double helix. For example, an average-sized human chromosome contains a single DNA double helix that is about 4.8 cm long with 140 million nucleotides on each strand; that is, 140 million base pairs (140 megabases (Mb)) of DNA.

To appreciate the difficulty in dealing with molecules this long in a cell only about 10 μm across, imagine a model of a human cell 1 meter across (a 10^5-fold increase in diameter). Now imagine the problem of fitting into this 1-meter-wide cell 46 DNA double helices that when scaled up by the same factor would each be just 0.2 mm thick but on average 4.8 km (about three miles) long. Then there is the challenge of replicating each of the DNA molecules and arranging for the cell to divide in such a way that the replicated DNA molecules are segregated equally into the two daughter cells. All this must be done in a way that avoids any tangling of the long DNA molecules.

To manage nuclear DNA molecules efficiently and avoid any tangling, they are complexed with various proteins and sometimes noncoding structural RNAs to form **chromatin** that undergoes different levels of coiling and compaction to form chromosomes. In interphase—the stages of the cell cycle other than mitosis (see Section 1.3)—the nuclear DNA molecules are still in a very highly extended form and normally the very long slender interphase chromosomes remain invisible under the light microscope. But even in interphase cells, the 2 nm-thick double helix is subject to at least two levels of coiling. First, the double helix is periodically wound round a specialized complex of positively charged histone proteins to form a 10 nm nucleosome filament. The nucleosome filament is then coiled into a 30 nm chromatin fiber that undergoes looping and is supported by a scaffold of nonhistone proteins (**Figure 1.7**).

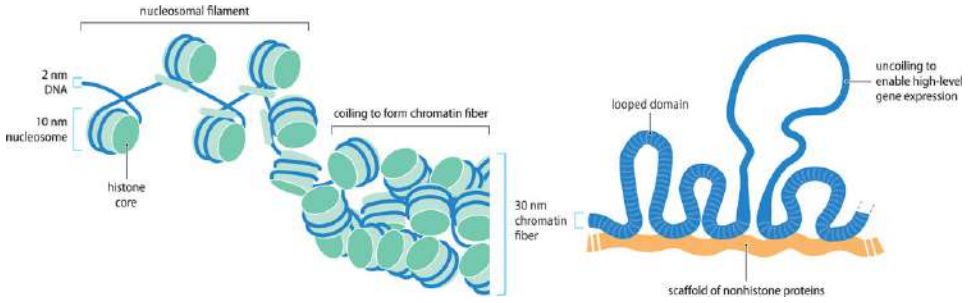


Figure 1.7 From DNA double helix to interphase chromatin. Binding of basic histone proteins causes the 2 nm DNA double helix to undergo coiling, forming first a 10 nm filament studded with nucleosomes that is further coiled to give a 30 nm chromatin fiber. In interphase, the chromatin fiber is organized in looped domains, each containing about 50–200 kilobases of DNA, that are attached to a central scaffold of nonhistone proteins. High levels of gene expression require local uncoiling of the chromatin fiber to give the 10 nm nucleosomal filaments. The diagram does not show structural RNAs that can be important in chromatin. (Adapted from Grunstein M [1992] *SciAm* 267:68–74; PMID 1411455. With permission from Macmillan Publishers Ltd; and Alberts B, Johnson A, Lewis J et al. [2008] Molecular Biology of the Cell, 5th ed. Garland Science.)

During interphase most chromatin exists in an extended state (**euchromatin**) that is dispersed through the nucleus. Euchromatin is not uniform, however—some euchromatic regions are more condensed than others, and genes may or may not be expressed, depending on the cell type and its functional requirements. Some chromatin, however, remains highly condensed throughout the cell cycle and is generally genetically inactive (**heterochromatin**).

As cells prepare to divide, the chromosomes need to be compacted much further to maximize the chances of correct pairing and segregation of chromosomes into daughter cells. Packaging of DNA into **nucleosomes** and then the 30 nm chromatin fiber results in a linear condensation of about 50-fold. During the M (mitosis) phase, higher-order coiling occurs (see Figure 1.7), so that DNA in a human metaphase chromosome is compacted to about 1/10 000 of its stretched-out length. As a result, the short, stubby metaphase chromosomes are readily visible under light microscopes.

Chromosome function: replication origins, centromeres, and telomeres

The DNA within a chromosome contains genes that are expressed according to the needs of a cell. But it also contains specialized sequences that are needed for chromosome function. Three major classes are described below.

Centromeres

When a cell divides the chromosomes must be correctly segregated to the two daughter cells. This requires a **centromere**, a region to which a pair of large protein complexes called kinetochores will bind just before the preparation for cell division (Figure 1.8). Centromeres can be seen at metaphase as the primary constriction that separates the short and long chromosome arms. Microtubules attached to each kinetochore are responsible for positioning the chromosomes correctly at metaphase and then pulling the separated chromosomes to opposite poles of the mitotic spindle.

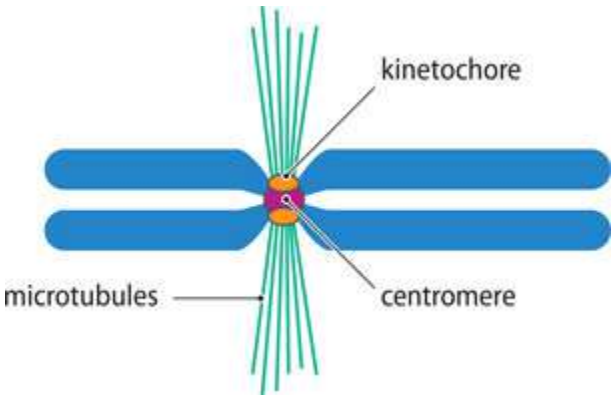


Figure 1.8 Centromere function relies on the assembly of kinetochores and attached microtubules.

The DNA sequences at centromeres are very different in different organisms. In a mammalian chromosome, the centromeric DNA is a

heterochromatic region dominated by highly repetitive DNA sequences that often extend over megabases of DNA.

Replication origins

For a chromosome to be replicated, it needs one or more replication origins —DNA sequence components to which protein factors bind in preparation for initiating DNA replication. The chromosomes of budding yeast can be replicated using a single very short highly defined DNA sequence, but in the cells of complex organisms, such as mammals, DNA is replicated at multiple initiation sites along each chromosome; the replication origins are quite long and do not have a common base sequence.

Telomeres

Telomeres are specialized structures at the ends of chromosomes that are necessary for the maintenance of chromosome integrity (if a telomere is lost after chromosome breakage, the resulting chromosome end is unstable; it tends to fuse with the ends of other broken chromosomes, or to be involved in recombination events, or to be degraded).

Unlike centromeric DNA, telomeric DNA has been well conserved during evolution. In vertebrates, the DNA of telomeres consists of many tandem (sequential) copies of the sequence TTAGGG to which certain telomeric proteins bind. Most of the telomere DNA is double-stranded with one strand containing TTAGGG repeats (the G-rich strand) and the complementary strand containing CCCTAA repeats (the C-rich strand). However, at its 3¢ end, the G-rich strand has an overhang (with about 30 TTAGGG repeats) that folds back and base pairs with the C-rich strand. The resulting T-loop is thought to protect the telomere DNA from natural cellular exonucleases that repair double-strand DNA breaks (**Figure 1.9**).

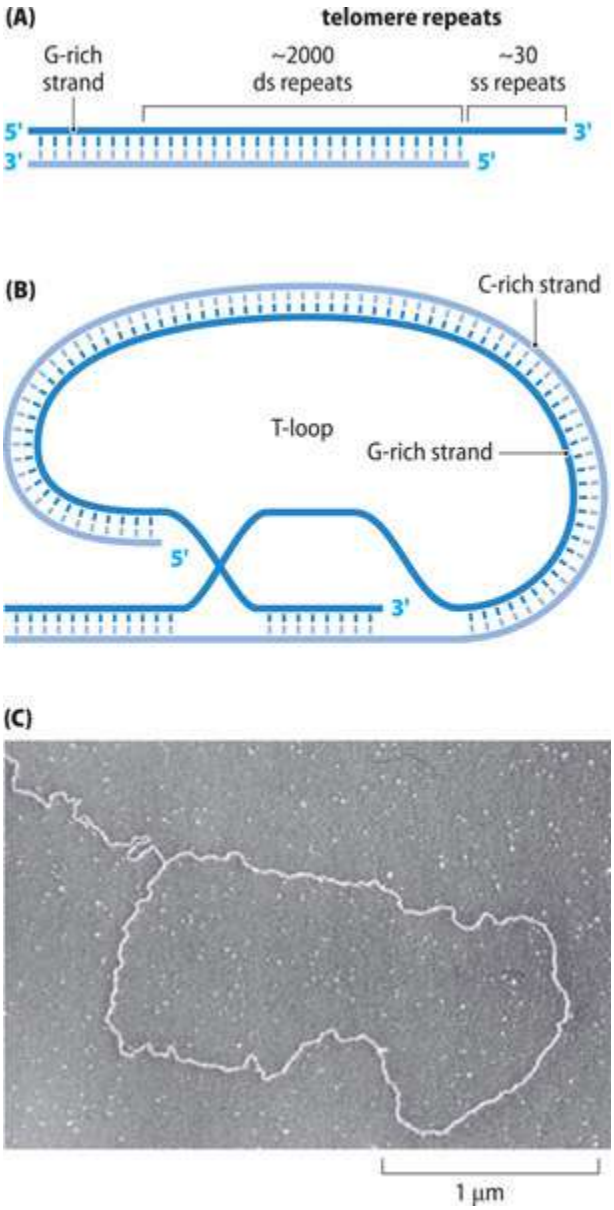


Figure 1.9 Telomere structure and T-loop formation. (A) Human telomere structure. A tandem array of roughly 2000 copies of the double-stranded hexanucleotide (TTAGGG/CCCTAA) repeat followed by a protrusion of about 30 single-stranded TTAGGG repeats. Abbreviations: ss, single-strand; ds, double-strand. (B) T-loop formation. The single-stranded terminus can loop back and invade the double-stranded region by base pairing with the complementary C-rich strand. (C) Electron micrograph showing formation of a roughly 15-kilobase T-loop at the end of an interphase human chromosome. (From Griffith JD et al. [1999] *Cell* 97:503–514; PMID 10338214. With permission from Elsevier.)

1.3 DNA AND CHROMOSOMES IN CELL DIVISION AND THE CELL CYCLE

Differences in DNA copy number between cells

Like other multicellular organisms, we have cells that are structurally and functionally diverse. In each individual the different cell types have the same genetic information, but only a subset of genes is expressed in each cell. What determines the identity of a cell—whether a cell is a B lymphocyte or a hepatocyte, for example—is the pattern of expression of the different genes across the genome.

As well as differences in gene expression, different cells can vary in the number of copies of each DNA molecule. The term *ploidy* describes the number of copies (n) of the basic chromosome set (the collective term for the different chromosomes in a cell) and also describes the copy number of each of the different nuclear DNA molecules.

The DNA content of a single chromosome set is represented as C. Human cells—and the cells of other mammals—are mostly **diploid** (2C), with nuclei containing two copies of each type of chromosome, one paternally inherited and one maternally inherited. Sperm and egg cells are **haploid** cells that contain only one of each kind of chromosome (1C). Human sperm and eggs each have 23 different types of chromosomes and so $n = 23$ in humans.

Some specialized human cells are nulliploid (0C) because they lack a nucleus—examples include erythrocytes, platelets, and terminally differentiated keratinocytes. Others are naturally polyploid (more than 2C). Polyploidy can occur by two mechanisms. The DNA might undergo multiple rounds of replication without cell division, as when the large megakaryocytes in blood are formed (they have from 16 to 64 copies of each chromosome, and the nucleus is large and multilobed). Alternatively, polyploid cells originate by cell fusion to give cells with multiple nuclei, as in the case of muscle fiber cells.

Mitochondrial DNA copy number

The great majority of our cells are diploid and contain two copies of each nuclear DNA molecule. In stark contrast, the number of copies of the mitochondrial DNA (mtDNA) can vary from hundreds to many thousands according to the cell type, and can even vary over time in some cells. The two types of haploid cells show very large differences in mtDNA copy number: a human sperm typically has about 100 mtDNA copies, but a human egg cell usually has about 250 000 mtDNA molecules.

The cell cycle and segregation of replicated chromosomes and DNA molecules

Cells also differ according to whether they actively participate in the cell cycle and undergo successive rounds of cell division. Each time a cell divides, it gives rise to two daughter cells. To keep the number of chromosomes constant there needs to be a tight regulation of chromosome replication and chromosome segregation. Each chromosome needs to be replicated just once to give rise to two daughter chromosomes, which must then segregate equally so that one passes to each daughter cell.

During normal periods of growth there is a need to expand cell number. In the fully grown adult, the majority of cells are terminally differentiated and do not divide, but stem cells and progenitor cells continue to divide to replace cells that have a high turnover, notably blood, skin, sperm, and intestinal epithelial cells.

Each round of the cell cycle involves a phase in which the DNA replicates—S phase (synthesis of DNA)—and a phase where the cell divides—M phase. Note that M phase involves both nuclear division (mitosis) and cell division (cytokinesis). In the intervals between these two phases are two gap phases—G_1 phase (gap between M phase and S phase) and G_2 phase (gap between S phase and M phase)—see **Figure 1.10**.

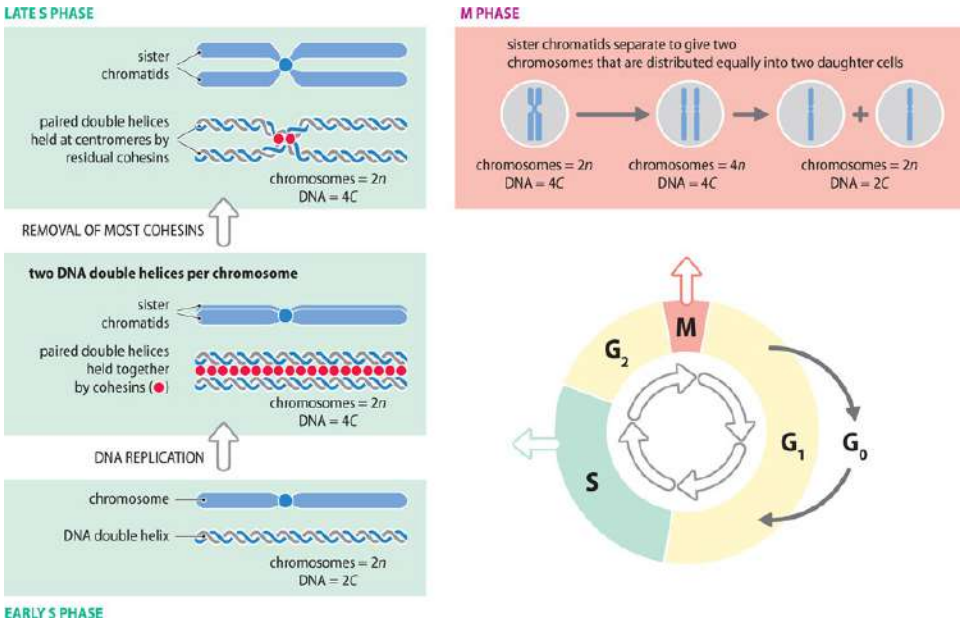


Figure 1.10 Changes in chromosomes and DNA content during the cell cycle. The cell cycle consists of four major phases as shown at the bottom right (in the additional G0 phase a cell exits from the cell cycle and remains suspended in a stationary phase that resembles G1 but can subsequently rejoin the cell cycle under certain conditions). In the expanded panels for M and S phases we show for convenience just a single chromosome, and we illustrate in each of the three boxes at left, representing S phase at different stages, how a single chromosome (top) relates to its DNA molecule (bottom). Chromosomes contain one DNA double helix from the end of M phase right through until just before the DNA duplicates in S phase. After duplication, the two double helices are held tightly together along their lengths by binding proteins called cohesins (red circles), and the chromosome now consists of two sister chromatids each having a DNA double helix. The sister chromatids becomes more obvious in late S phase when most of the cohesins are removed except for some at the centromere, which continue to hold the two sister chromatids together. The sister chromatids finally separate in M phase to form two independent chromosomes that are then segregated into the daughter cells. Note that the S-phase chromosomes in the boxes at left are shown, purely for convenience, in a compact form; in reality they are enormously extended.

Cell division takes up only a brief part of the cell cycle. For actively dividing human cells, a single turn of the cell cycle might take about 24 hours; M phase often occupies about 1 hour. During the short M phase, the

chromosomes become extremely highly condensed in preparation for nuclear and cell division. After M phase, cells enter a long growth period called **interphase** (= G_1 + S + G_2 phases), during which chromosomes are enormously extended, allowing genes to be expressed.

G_1 is the long-term end state of terminally differentiated nondividing cells. For dividing cells, the cells will enter S phase only if they are committed to mitosis; if not, they are induced to leave the cell cycle to enter a resting phase, the G_0 phase (a modified G_1 stage). When conditions become suitable later on, cells may subsequently move from G_0 to re-enter the cell cycle.

Changes in cell chromosome number and DNA content

During the cell cycle, the amount of DNA in a cell and the number of chromosomes change. In the box panels in Figure 1.10 we follow the fate of a single chromosome through M phase and then through S phase. If we were to consider a diploid human cell this would be one chromosome out of the 46 ($2n$) chromosomes present after daughter cells are first formed. We also show in this figure how a single chromosome (top) relates to its DNA double helix content at different stages in S phase. The progressive changes in the number of chromosomes and the DNA content of cells at different stages of the cell cycle are listed below.

- From the end of the M phase right through until DNA duplication in S phase, each chromosome of a diploid ($2n$) cell contains a single DNA double helix; the total DNA content is therefore 2C.
- After DNA duplication, the total DNA content per cell is 4C, but specialized binding proteins called cohesins hold the duplicated double helices together as **sister chromatids***within* a single chromosome. The chromo-some number remains the same ($2n$), but each chromosome now has double the DNA content of a chromosome in early S phase. In late S phase, most of the cohesins

are removed but cohesins at the centromere are retained to keep the sister chromatids together.

- During M phase, the residual cohesins are removed and the duplicated double helices finally separate. That allows sister chromatids to separate to form two daughter chromosomes, giving 4*n* chromosomes. The duplicated chromosomes segregate equally to the two daughter cells so that each will have 2*n* chromosomes and a DNA content of 2C.

Figure 1.10 can give the misleading impression that all the interesting action happens in S and M phases. That is quite wrong—a cell spends most of its life in the G_0 or G_1 phases, and that is where the genome does most of its work, issuing the required instructions to make the diverse protein and RNA products needed for cells to function.

Mitochondrial DNA replication and segregation

In advance of cell division, mitochondria increase in mass and mtDNA molecules replicate before being segregated into daughter mitochondria that then need to segregate into daughter cells. Whereas the replication of nuclear DNA molecules is tightly controlled, the replication of mtDNA molecules is not directly linked to the cell cycle.

Replication of mtDNA molecules simply involves increasing the number of DNA copies in the cell, without requiring equal replication of individual mtDNAs. That can mean that some individual mtDNAs might not be replicated and other mtDNA molecules might be replicated several times (**Figure 1.11**).

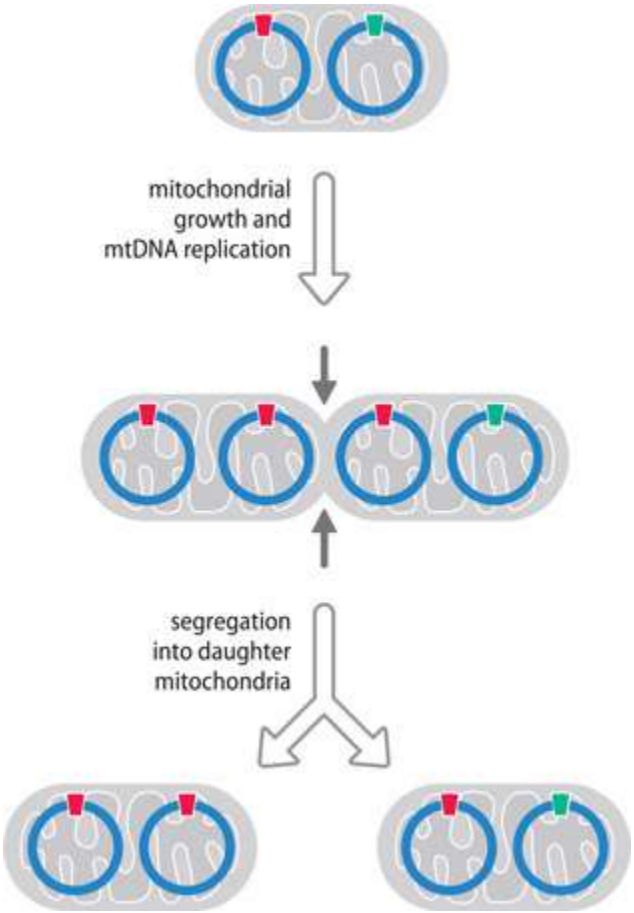


Figure 1.11 Unequal replication of individual mitochondrial DNAs. Unlike in the nucleus, where replication of a chromosomal DNA molecule is tightly controlled and normally produces two copies, mitochondrial DNA (mtDNA) replication is stochastic. When a mitochondrion increases in mass in preparation for cell division, the overall amount of mitochondrial DNA increases in proportion, but individual mtDNAs replicate unequally. In this example, the mtDNA with the green tag fails to replicate and the one with the red tag replicates to give three copies. Variants of mtDNA can arise through mutation so that a person can inherit a mixed population of mtDNAs (heteroplasmy). Unequal replication of pathogenic and nonpathogenic mtDNA variants can have important consequences, as described in Chapter 5.

Whereas the segregation of nuclear DNA molecules into daughter cells needs to be equal and is tightly controlled, segregation of mtDNA molecules into daughter cells can be unequal. Even if the segregation of mtDNA molecules into daughter mitochondria is equal (as shown in Figure

1.11), the segregation of the mitochondria into daughter cells is thought to be stochastic.

Mitosis: the usual form of cell division

Most cells divide by a process known as mitosis. In the human life cycle, mitosis is used to generate extra cells that are needed for periods of growth and to replace various types of short-lived cells. Mitosis ensures that a single parent cell gives rise to two daughter cells that are both genetically identical to the parent cell (barring any errors that might have occurred during DNA replication). During a human lifetime, there may be something like 10^{17} mitotic divisions.

The M phase of the cell cycle includes both nuclear division (mitosis, which is divided into the stages of prophase, prometaphase, metaphase, anaphase, and telophase), and also cell division (cytokinesis), which overlaps the final stages of mitosis (**Figure 1.12**). In preparation for cell division, the previously highly extended duplicated chromosomes contract and condense so that, by the metaphase stage of mitosis, they are readily visible when viewed under the microscope.

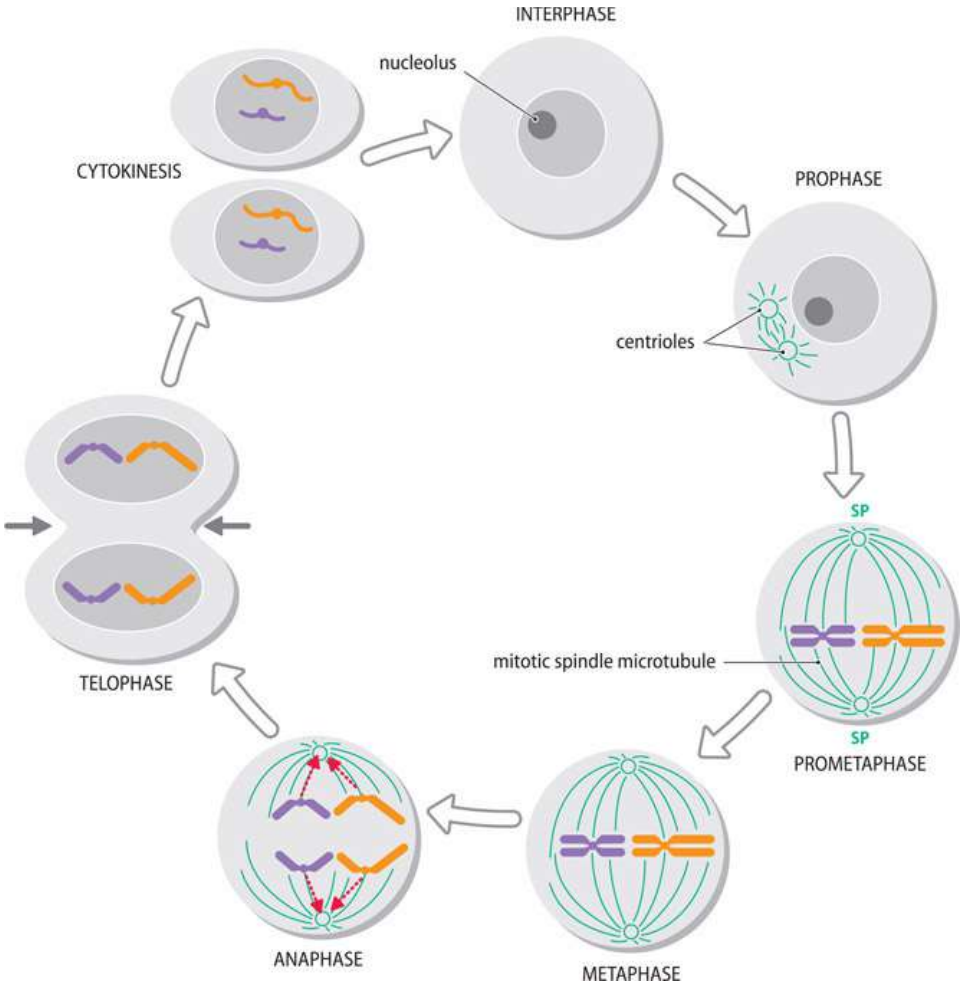


Figure 1.12 Mitosis (nuclear division) and cytokinesis (cell division). Early in prophase, centrioles (short cylindrical structures made up of microtubules and associated proteins) begin to separate and migrate to opposite poles of the cell. They give rise to the spindle poles (SP) from which microtubules will extend to the center of the cell to form the mitotic spindle. In prometaphase, the nuclear envelope breaks down, and the now highly condensed chromosomes become attached at their centromeres to the array of mitotic spindle microtubules. At metaphase, the chromosomes all lie along the middle of the mitotic spindle, still with the sister chromatids bound together (because of residual cohesins at the centromere that hold the duplicated DNA helices together). Removal of the residual cohesins allows the onset of anaphase: the sister chromatids separate and begin to migrate toward opposite poles of the cell (shown by dashed red arrows). The nuclear envelope forms again around the daughter nuclei during telophase, and the chromosomes decondense, completing mitosis. Before the final stages of mitosis, and most obviously at telophase, cytokinesis begins. The cell

becomes progressively constricted at its middle (shown at telophase by converging horizontal arrows), eventually resulting in full cytokinesis to produce two daughter cells.

The chromosomes of early S phase have one DNA double helix; however, after DNA replication, two identical DNA double helices are produced and held together by cohesins. Later, when the chromosomes undergo compaction in preparation for cell division, the cohesins are removed from all parts of the chromosomes apart from the centromeres. As a result, as early as prometaphase (when the chromosomes are now visible under the light microscope), individual chromosomes can be seen to comprise two **sister chromatids** that remain attached at the centromere (bound by some residual cohesins).

Later, at the start of anaphase, the remaining cohesins are removed and the two sister chromatids can now disengage to become independent chromosomes that will be pulled to opposite poles of the cell and then distributed equally to the daughter cells (see Figure 1.12).

Meiosis: a specialized reductive cell division giving rise to sperm and egg cells

The **germ line** is the collective term for cells that can pass genetic material to the next generation. It includes haploid sperm and egg cells (the **gametes**) and all the diploid precursor cells from which they arise by cell division, going all the way back to the zygote. The nongermline cells are known as **somatic cells**.

In humans, where $n = 23$, each gamete contains one sex chromosome plus 22 nonsex chromosomes (**autosomes**). In eggs the sex chromosome is always an X; in sperm it may be either an X or a Y. After a haploid sperm fertilizes a haploid egg, the resulting diploid **zygote** and almost all of its descendant cells have the chromosome constitution 46,XX (female) or 46,XY (male).

Diploid primordial germ cells migrate into the embryonic gonad and engage in repeated rounds of mitosis, to generate spermatogonia in males

and oogonia in females. Further growth and differentiation produce primary spermatocytes in the testis and primary oocytes in the ovary. The diploid spermatocytes and oocytes can then undergo **meiosis**, the cell division process that produces haploid gametes.

Meiosis is a *reductive* division because it involves two successive cell divisions (meiosis I and meiosis II) but only one round of DNA replication (**Figures 1.13** and **1.14**). As a result, it gives rise to four haploid cells. In males, the two meiotic cell divisions are each symmetric, producing four functionally equivalent spermatozoa. Huge numbers of sperm are produced, and spermatogenesis is a continuous process from puberty onward.

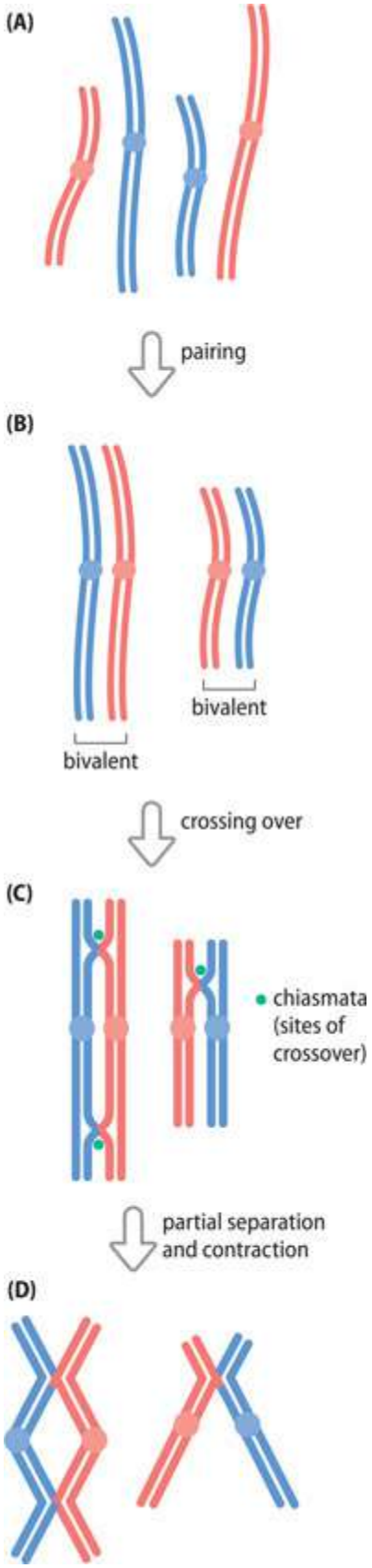
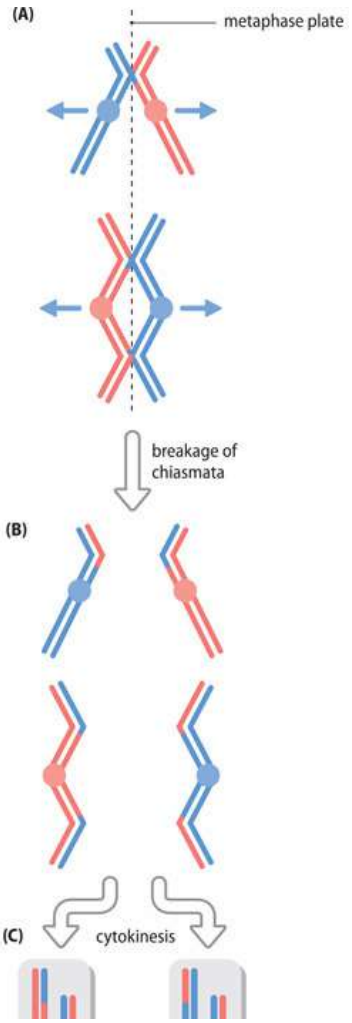

(A)
pairing
(B)
bivalent
bivalent
crossing over
(C)
chiasmata (sites of crossover)
partial separation and contraction
(D)

Figure 1.13 Prophase stages in meiosis I. (A) In leptotene, the duplicated chromosomes (each with a pair of sister chromatids) begin to condense but remain unpaired. (B) In zygotene, pairing of maternal and paternal homologous chromosomes (**homologs**) occurs, to form bivalents with four chromatids. (C) In pachytene, recombination (crossing over) occurs through the physical breakage and subsequent rejoining of maternal and paternal chromosome fragments. There are two chiasmata (crossovers) in the bivalent on the left, and one in the bivalent on the right. For simplicity, both chiasmata on the left involve the same two chromatids. In reality, more chiasmata may occur, involving three or even all four chromatids in a bivalent. (D) During diplotene, the homologous chromosomes may separate slightly, except at the chiasmata. A further stage, diakinesis, is marked by contraction of the bivalents and is the transition to metaphase I. In this figure, only 2 of 23 possible pairs of homologs are illustrated (with the maternal homolog colored pink, and the paternal homolog blue).



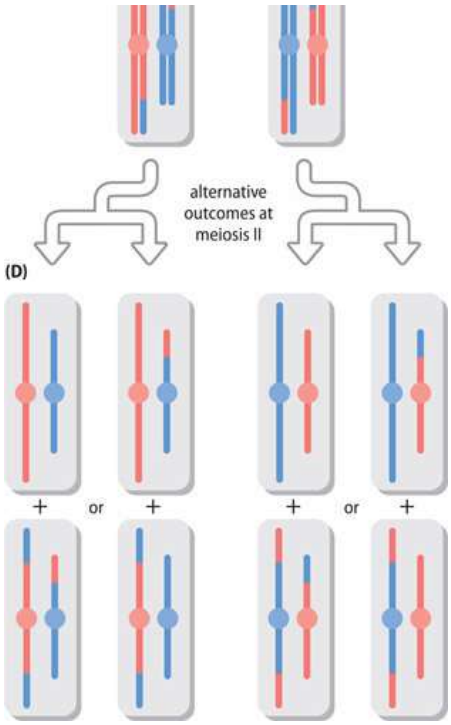


Figure 1.14 Metaphase I to production of gametes. (A) At metaphase I, the bivalents align on the metaphase plate, at the center of the spindle apparatus. Contraction of spindle fibers draws the chromosomes in the direction of the spindle poles (arrows). (B) The transition to anaphase I occurs at the consequent rupture of the chiasmata. (C) Cytokinesis segregates the two chromosome sets, each to a different primary spermatocyte. Note that, after recombination during prophase I (see Figure 1.13C), the chromatids share a single centromere but are no longer identical. (D) Meiosis II in each primary spermatocyte, which does not include DNA replication, generates unique genetic combinations in the haploid secondary spermatocytes. Only 2 of the possible 23 different human chromosomes are depicted, for clarity, so only 2^2 (that is, 4) of the possible 2^{23} (8 388 608) possible combinations are illustrated. Although oogenesis can produce only one functional haploid gamete per meiotic division, the processes by which genetic diversity arises are the same as in spermatogenesis.

Female meiosis is different: cell division is asymmetric, resulting in unequal division of the cytoplasm. The products of female meiosis I (the first meiotic cell division) are a large secondary oocyte and a small cell, the *polar body,* which is discarded. During meiosis II the secondary oocyte then

gives rise to the large mature egg cell and a second polar body (which again is discarded).

In humans, primary oocytes enter meiosis I during fetal development but are then all arrested at prophase until after the onset of puberty. After puberty in females, one primary oocyte completes meiosis with each menstrual cycle. Because ovulation can continue up to the fifth and sometimes sixth decades, this means that meiosis can be arrested for many decades in those primary oocytes that are not used in ovulation until late in life.

Pairing of paternal and maternal homologs (synapsis)

Each of our diploid cells contains two copies (**homologs**) of each type of chromosome, one maternal copy and one paternal copy. So, for example, paternal chromosome 1 and maternal chromosome 1 are homologs. The exception, of course, are the X and Y chromosomes in males.

A special feature of meiosis I—that distinguishes it from mitosis and meiosis II—is the pairing *(synapsis)* of paternal and maternal homologs. Then, the maternal and paternal homologs, each with sister chromatids following DNA replication, align along their lengths and become bound together. The resulting *bivalent* has four strands: two paternally inherited sister chromatids and two maternally inherited sister chromatids (see Figures 1.13B–D and 1.14).

The pairing of homologs is required for recombination to occur (as described in the next subsection). It must ultimately be dictated by high levels of DNA sequence identity between the homologs. The high sequence matching between homologs required for pairing does not need to be complete, however: when there is some kind of chromosome abnormality so that the homologs do not completely match, the matching segments usually manage to pair up.

Pairing of maternal and paternal sex chromosomes is straightforward in female meiosis, but in male meiosis there is the challenge of pairing a

maternally inherited X chromosome with a paternally inherited Y. The human X chromosome is very much larger than the Y, and their DNA sequences are very different. However, they do have some sequences in common, notably a major *pseudoautosomal region* located close to the short-arm telomeres. The X and Y chromosomes cannot pair up along their lengths, but because they have some sequences in common, they can always pair up along these regions. We will explore this in greater detail in Chapter 5 when we consider pseudoautosomal inheritance.

Recombination

The prophase of meiosis I begins during fetal life and, in human females, can last for decades. During this extended process, parternal and maternal chromatids within each bivalent normally exchange segments of DNA at randomly positioned but matching locations. This process—called **recombination** (or crossover)—involves physical breakage of the DNA in one paternal and one maternal chromatid, and the subsequent joining of maternal and paternal fragments.

Recombined homologs seem to be physically connected at specific points. Each such connection marks the point of crossover and is known as a chiasma (plural chiasmata—see Figure 1.13C). The distribution of chiasmata across chromosomes is nonrandom. The number of chiasmata per meiosis shows significant sex differences, and there are very significant differences between individuals of the same sex (and even between individual meioses from a single individual). In a large recent study of human meiosis, an average of 38 recombinations were detected per female meiosis, while 24 meioses occurred on average in male meiosis but with very significant variation (shown in Figure 8.3 on page 244). In addition to their role in recombination, chiasmata are thought to be essential for correct chromosome segregation during meiosis I.

There are hotspot regions where recombination is more likely to occur. For example, recombination is more common in subtelomeric regions. In

the case of X–Y crossover there is an obligate crossover within a short 2.6 Mb pseudoautosomal region located at the tips of the short arms of the X and Y. This region is so called because it is regularly swapped between the X and Y chromosomes and so the inheritance pattern for any DNA variant here is not X-linked or Y-linked but instead resembles autosomal inheritance.

Why each of our gametes is unique

The sole purpose of sex in biology is to produce novel combinations of gene variants, and the instrument for achieving that aim is meiosis. The whole point of meiosis is to produce *genetically unique* gametes by selecting different combinations of DNA sequences on maternal and paternal homologs.

Although a single ejaculate may contain hundreds of millions of sperm, meiosis ensures that no two sperm will be genetically identical. Equally, no two eggs are genetically identical. Each zygote must also be unique because at fertilization a unique sperm combines with a unique egg. However, a unique fertilization event can occasionally give rise to two genetically identical (**monozygotic**) twins if the embryo divides into two at a very early stage in development (monozygotic twins are nevertheless unique individuals—genetics is not everything in life!).

The second division of meiosis is identical in form to mitosis; meiosis I is where the genetic diversity originates, and that involves two mechanisms. First, there is independent assortment of paternal and maternal homologs. After DNA replication, the homologous chromosomes each comprise two sister chromatids, so each bivalent is a four-stranded structure at the metaphase plate. Spindle fibers then pull one complete chromosome (two chromatids) to either pole. In humans, for each of the 23 homologous pairs, the choice of which daughter cell each homolog will enter is independent. This allows 2^{23} or about 8.4×10^6 different possible combinations of parental chromosomes in the gametes that might arise from a single meiotic division (**Figure 1.15**).

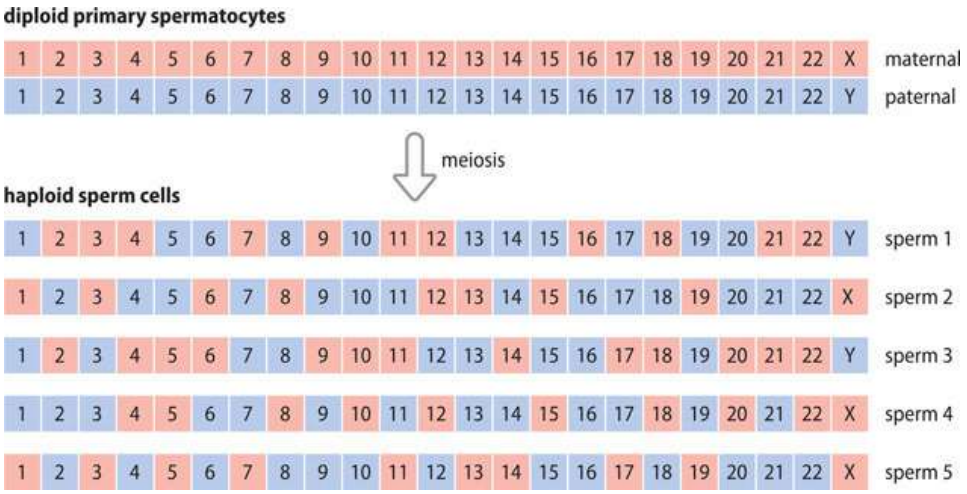


Figure 1.15 Independent assortment of maternal and paternal homologs during meiosis. The figure shows a random selection of just 5 of the 8 388 608 (2^{23}) theoretically possible combinations of homologs that might occur in haploid human spermatozoa after meiosis in a diploid primary spermatocyte. Maternally derived homologs are represented by pink boxes, and paternally derived homologs by blue boxes. For simplicity, the diagram ignores recombination—but see Figure 1.16.

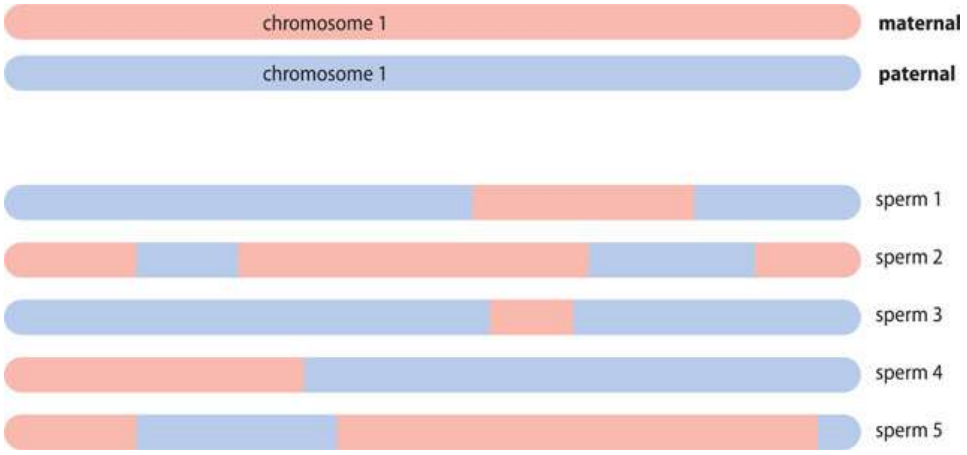


Figure 1.16 Recombination superimposes additional genetic variation at meiosis I. Figure 1.15 illustrates the contribution to genetic variation at meiosis I made by independent assortment of homologs, but for simplicity it ignores the contribution made by recombination. In reality each transmitted chromosome is a mosaic of paternal and maternal DNA sequences, as shown here. See Figure 8.1 on page 243 for a real-life example.

The second mechanism that contributes to genetic diversity is recombination. Whereas sister chromatids within a bivalent are genetically identical, the paternal and maternal chromatids are not. On average their DNA will differ at roughly 1 in every 1000 nucleotides. Swapping maternal and paternal sequences by recombination will therefore produce an extra level of genetic diversity (**Figure 1.16**). It raises the number of permutations from the 8.4 million that are possible just from the independent assortment of maternal and paternal homologs alone, to a virtually infinite number.

SUMMARY

- Nucleic acids are negatively charged long polymers composed of sequential nucleotides that each consist of a sugar, a nitrogenous base, and a phosphate group. They have a sugar-phosphate backbone with bases projecting from the sugars.
- Nucleic acids have four types of bases: adenine (A), cytosine (C), guanine (G), and either thymine (T) in DNA or uracil (U) in RNA. The sequence of bases determines the identity of a nucleic acid and its function.
- RNA normally consists of a single nucleic acid chain, but in cells a DNA molecule has two chains (strands) that form a stable duplex (in the form of a double helix). Duplex formation requires hydrogen bonding between matched bases (base pairs) on the two strands.
- In DNA two types of base pairing exist: A pairs with T, and C pairs with G. According to these rules, the two strands of a DNA double helix are said to have complementary base sequences.
- Base pairing also occurs in RNA and includes G–U base pairs, as well as G–C and A–U base pairs. Two different RNA

molecules with partly complementary sequences can associate by forming hydrogen bonds. Intramolecular hydrogen bonding also allows a single RNA chain to form a complex three-dimensional structure.

- DNA carries primary instructions that determine how cells work and how an individual is formed. Defined segments of DNA called genes are used to make a single-stranded RNA copy that is complementary in sequence to one of the DNA strands (transcription).
- DNA is propagated from one cell to daughter cells by replicating itself. The two strands of the double helix are unwound, and each strand is used to make a new complementary DNA copy. The two new nuclear DNA double helices (each with one parental DNA strand and one new DNA strand) are segregated so that each daughter cell receives one DNA double helix.
- RNA molecules function in cells either as a mature noncoding RNA, or as a messenger RNA with a coding sequence used to make the polypeptide chain of a protein (translation).
- Each nuclear DNA molecule is complexed with different proteins and some noncoding RNAs to form a chromosome that condenses the DNA and protects it.
- Packaging DNA into chromosomes stops the long DNA chains from getting entangled within cells, and by greatly condensing the DNA in preparation for cell division it allows the DNA to be segregated correctly to daughter cells and to offspring.
- Our sperm and egg cells are haploid cells with a set of 23 different chromosomes (each with a single distinctive DNA molecule). There is one sex chromosome (an X chromosome in eggs; either an X or Y in sperm) and 22 different autosomes (nonsex chromosomes).

- Most of our cells are diploid with two copies of the haploid chromosome set, one set inherited from the mother and one from the father. Maternal and paternal copies of the same chromosome are known as homologs.
- There is one type of mitochondrial DNA (mtDNA); it is present in many copies with wide variation in copy number between different cell types. Both the replication of mtDNA and its segregation to daughter cells occur stochastically.
- Cells need to divide as we grow. In fully formed adults, most of our cells are specialized, nondividing cells, but some cells are required to keep on dividing to replace short-lived cells, such as blood, skin, and intestinal epithelial cells.
- Mitosis is the normal form of cell division. Each chromosome (and chromosomal DNA) replicates once and the duplicated chromosomes are segregated equally into the two daughter cells.
- Meiosis is a specialized form of cell division required to produce haploid sperm and egg cells. The chromosomes in a diploid spermatogonium or oogonium replicate once, but there are two successive cell divisions to reduce the number of chromosomes in each cell.
- Each sperm cell produced by a man is unique, as is each egg cell that a woman produces. During the first cell division in meiosis, maternal and paternal homologs associate and exchange sequences by recombination. Largely random recombination results in unpredictable new DNA sequence combinations in each sperm and in each egg.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

More detailed treatment of the subject matter in this chapter can be found in more comprehensive genetics and cell biology textbooks such as:

Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K & Walter P (2015) *Molecular Biology of the Cell*, 6th ed. Garland Science.

Strachan T & Read AP (2019) *Human Molecular Genetics*, 5th ed. CRC Press, Taylor & Francis.

2

Fundamentals of gene structure, gene expression, and human genome organization

DOI: 10.1201/9781003044406-2

CONTENTS

Our genome is complex, comprising about 3.2 Gb (3.2×10^9 base pairs; Gb = giga-base) of DNA. One of its main tasks is to produce a huge variety of different proteins that dictate how our cells work. Surprisingly, however, **coding DNA**—DNA sequences that specify the polypeptides of our proteins—account for just about 1.5 % of our DNA.

The remainder of our genome is noncoding DNA that does not make any protein. A significant fraction of the noncoding DNA is functionally important, including many different classes of DNA regulatory sequences that control how our genes work (such as promoters and enhancers), and DNA sequences that specify short regulatory sequence elements that work at the RNA level.

Additionally, we have many thousands of genes that do not make polypep-tides; instead they make different classes of functional noncoding RNA. Some of these **RNA genes**—such as genes encoding ribosomal RNA and transfer RNA needed for protein synthesis—have been known for decades, but one of the big surprises in recent years has been the sheer number and variety of noncoding RNAs in our cells. In addition to the RNA genes, our protein-coding genes frequently make noncoding RNA transcripts as well as messenger RNAs (mRNAs).

Like other complex genomes, our genome has a large proportion of moderately to highly repetitive DNA sequences. Some of these are important in centro-mere and telomere function; others are

important in genome evolution.

By 2003, the Human Genome Project (HGP) provided the first comprehensive insights into our genome, delivering an essentially complete nucleotide sequence of the gene-rich euchromatic component of the genome. Follow-up studies have compared our genome with other genomes, helping us to understand how our genome evolved. The comparative genomics studies, together with genome-wide functional and bioinformatic analyses, are providing major insights into how our genome works.

2.1 PROTEIN-CODING GENES: STRUCTURE AND EXPRESSION

Proteins are the main functional endpoints of gene expression and perform a huge diversity of roles that govern how cells work (acting as structural components, enzymes, carrier proteins, ion channels, signaling molecules, gene regulators, and so on). They each consist of one or more polypeptides, long sequences of amino acids that are encoded by a coding DNA. In many cases a protein also contains carbohydrate or lipid components (which are not genetically determined).

Protein-coding genes come in a startling variety of organizations, as described below, and synthesize one or more polypeptides. Polypeptide synthesis is not the endpoint, however. A newly synthesized polypeptide must undergo multiple different maturation steps, usually involving chemical modification and cleavage events, and often then associates with other polypeptides to form a working protein.

Gene organization: exons and introns

The protein-coding genes of bacteria are small (on average about 1000 bp long) and simple. The gene is transcribed to give an mRNA with a continuous coding sequence that is then translated to give a linear sequence of about 300 amino acids on average. Unexpectedly, the genes of eukaryotes turned out to be much bigger and much more complex than anticipated. And, as we will see, our protein-coding genes often contain a rather small amount of coding DNA.

For most eukaryotic protein-coding genes, the coding DNA is split into segments (**exons)** separated by noncoding DNA sequences (**introns)**. The number of exons and introns in a gene varies considerably (there seems little logic about precisely where introns insert within genes).

Excluding single-exon genes (some genes lack introns), average exon lengths show moderate variation from gene to gene, but introns can show extraordinary size differences. Our genes are therefore often large, sometimes extending over more than a megabase of DNA (Table 2.1).

TABLE 2.1 EXAMPLES OF DIFFERENTIAL GENE ORGANIZATION FOR HUMAN PROTEIN-CODING GENES

Human gene	Size in genome (kb)	No. of exons	Average size of exon (bp) *	Average size of intron (bp) **
SRY (male sex-determinant)	0.9	1	850	-

Items in brackets show the protein name. kb, kilobases (= 1000 bp).

* Note that the shortest human exon is just two nucleotides long, and final exons can quite often be long, the record being 27 303 bp.

** The shortest human intron is26bp, and the longest is 1160 411 bp–seePMID31164174

Human gene	Size in genome (kb)	No. of exons	Average size of exon (bp) *	Average size of intron (bp) **
HBB (-globin)	1.6	3	150	490
TP53 (p53)	39	10	236	3076
F8 (factor VIII)	186	26	375	7100
CFTR (cystic fibrosis transmembrane regulator)	250	27	227	9100
DMD (dystrophin)	2400	79	180	30 770

Items in brackets show the protein name. kb, kilobases (= 1000 bp).

* Note that the shortest human exon is just two nucleotides long, and final exons can quite often be long, the record being 27 303 bp.

** The shortest human intron is26bp, and the longest is 1160 411 bp–seePMID31164174

RNA splicing: stitching together the genetic information in exons

Like all genes, genes that are split into exons are initially transcribed by an RNA polymerase to give a long RNA transcript. This primary transcript is identical in base sequence to the transcribed region of the sense DNA strand, except that U replaces T (the transcribed region of DNA is called a **transcription unit**). Thereafter, the primary RNA transcript undergoes a form of processing called **RNA splicing**.

RNA splicing involves first cleaving the RNA transcript at the junctions between transcribed exons and introns. The individual transcribed intron sequences are often degraded, but the transcribed exon sequences are then covalently linked (spliced) in turn to make a mature RNA (Figure 2.1). RNA splicing is performed within the nucleus by spliceosomes, complex assemblies of protein factors and small nuclear RNA (snRNA) molecules.

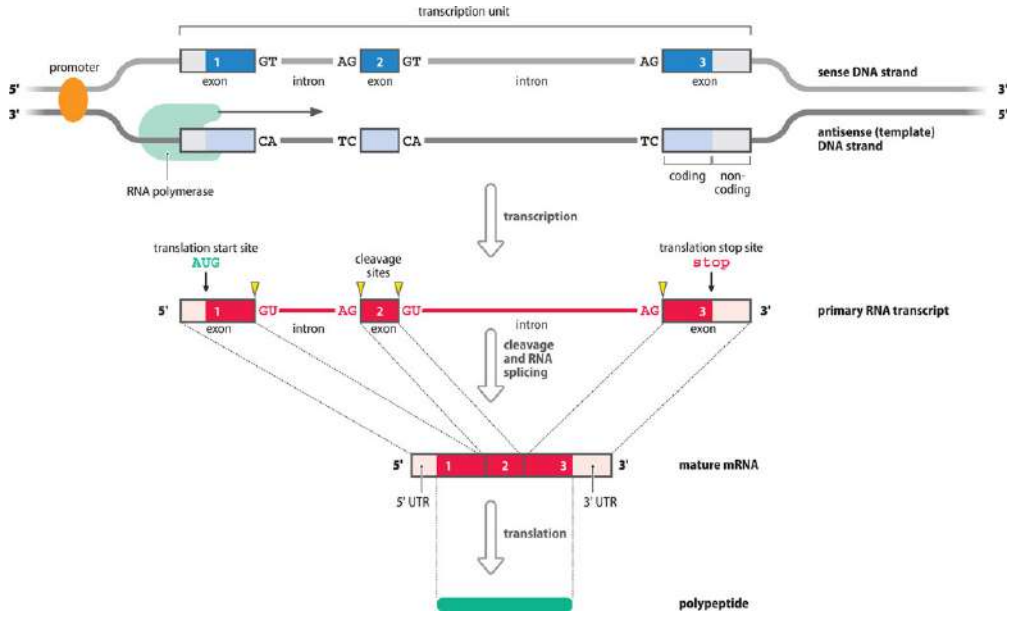


Figure 2.1 RNA splicing brings transcribed exon sequences together. Most of our protein-coding genes (and many RNA genes) undergo RNA splicing. In this generalized example a protein-coding gene is illustrated with an upstream promoter and three exons separated by two introns that each begin with the dinucleotide GT and end in the dinucleotide

AG. The central exon (exon 2) is composed entirely of coding DNA (deep red), but exons 1 and 3 have both coding DNA and also noncoding DNA sequences (shown in pink; they will eventually be used to make untranslated sequences in the mRNA). The three exons and the two separating introns are transcribed together to give a large primary RNA transcript. The RNA transcript is cleaved at positions corresponding to exonintron boundaries. The two transcribed intron sequences that are excised are each degraded, but the transcribed exon sequences are joined *(spliced)* together to form a contiguous mature RNA that has noncoding sequences at both the 5¢ and 3¢ ends. In the mature mRNA these terminal sequences will not be translated and so are known as *untranslated regions* (UTRs). The central coding sequence of the mRNA is defined by a translation start site (which is almost always the trinucleotide AUG) and a translation stop site, and is read *(translated)* to produce a polypeptide.

We do not fully understand how the spliceosome is able to recognize and cut the primary transcript at precise positions marking the start and end of introns. However, we do know that certain sequences are important in signaling the splice sites that define exon-intron boundaries. For example, very nearly all introns begin with a GT dinucleotide on the sense DNA strand and end with an AG so that the transcribed intron sequence begins with a GU (that marks the **splice donor site**) and ends in an AG (marking the **splice acceptor site**). The GT (GU) and AG end sequences need to be embedded in broader splice site consensus sequences that we will describe in Section 6.1 when we consider how gene expression is regulated. As we will see in Chapter 7, mutations at splice sites are important causes of disease.

Figure 2.1 might give the erroneous impression that all protein-coding genes undergo a specific, single type of RNA splicing. However, close to 10 % of our protein-coding genes have a single, uninterrupted exon and do not undergo RNA splicing at all—notable examples include histone genes. And most of the genes that go through RNA splicing undergo alternative RNA splicing patterns; a single gene can therefore produce different gene products that may be functionally different. We consider the concept of alternative splicing in greater detail in Chapter 6, in the context of gene regulation.

The evolutionary value of RNA splicing

As we will see in Section 2.2, many RNA genes also undergo RNA splicing. At this stage, one might reasonably wonder why RNA splicing is so important in eukaryotic cells, and so especially prevalent in complex multicellular organisms. Why do we need to split the genetic information in genes into sometimes so many different little exons? The answer is to help stimulate the formation of novel genes and novel gene products that can permit greater functional complexity during evolution.

The huge complexity of humans and other multicellular organisms has been driven by genome evolution. In addition to periodic gene duplication, various genetic mechanisms allow individual exons to be duplicated or swapped from one gene to another on an evolutionary timescale. That allows different ways of combining exons to produce novel hybrid genes. An additional source of complexity comes from using different combinations of exons to make alternative transcripts from the same gene (alternative splicing).

Translation: decoding messenger RNA to make a polypeptide

Messenger RNA (mRNA) molecules produced by RNA splicing in the nucleus are exported to the cytoplasm. Here they are bound by ribosomes, very large complexes consisting of four types of ribosomal RNA (rRNA) and many different proteins.

Although an mRNA is formed from exons only, it has sequences at its 5¢ and 3¢ ends that are noncoding. Having bound to mRNA, the job of the ribosomes is to scan the mRNA sequence to find and interpret a central coding sequence that will be translated to make a polypeptide. The noncoding sequences at the ends are known as **untranslated regions** (UTRs; as shown in Figure 2.1). and contain sequences that are important in regulating gene expression.

A polypeptide is a polymer made up of a linear sequence of **amino acids** (**Figure 2.2A**). Amino acids have the general formula NH2-CH(R)-COOH, where R is a variable side chain that defines the chemical identity of the amino acid and is connected to the central (alpha) carbon of the NH-CH-CO framework sequence. There are 20 common amino acids (Figure 2.2C). Polypeptides are made by a condensation reaction between the carboxyl (COOH) group of one amino acid and the amino (NH2) group of another amino acid, forming a peptide bond (see Figure 2.2B).

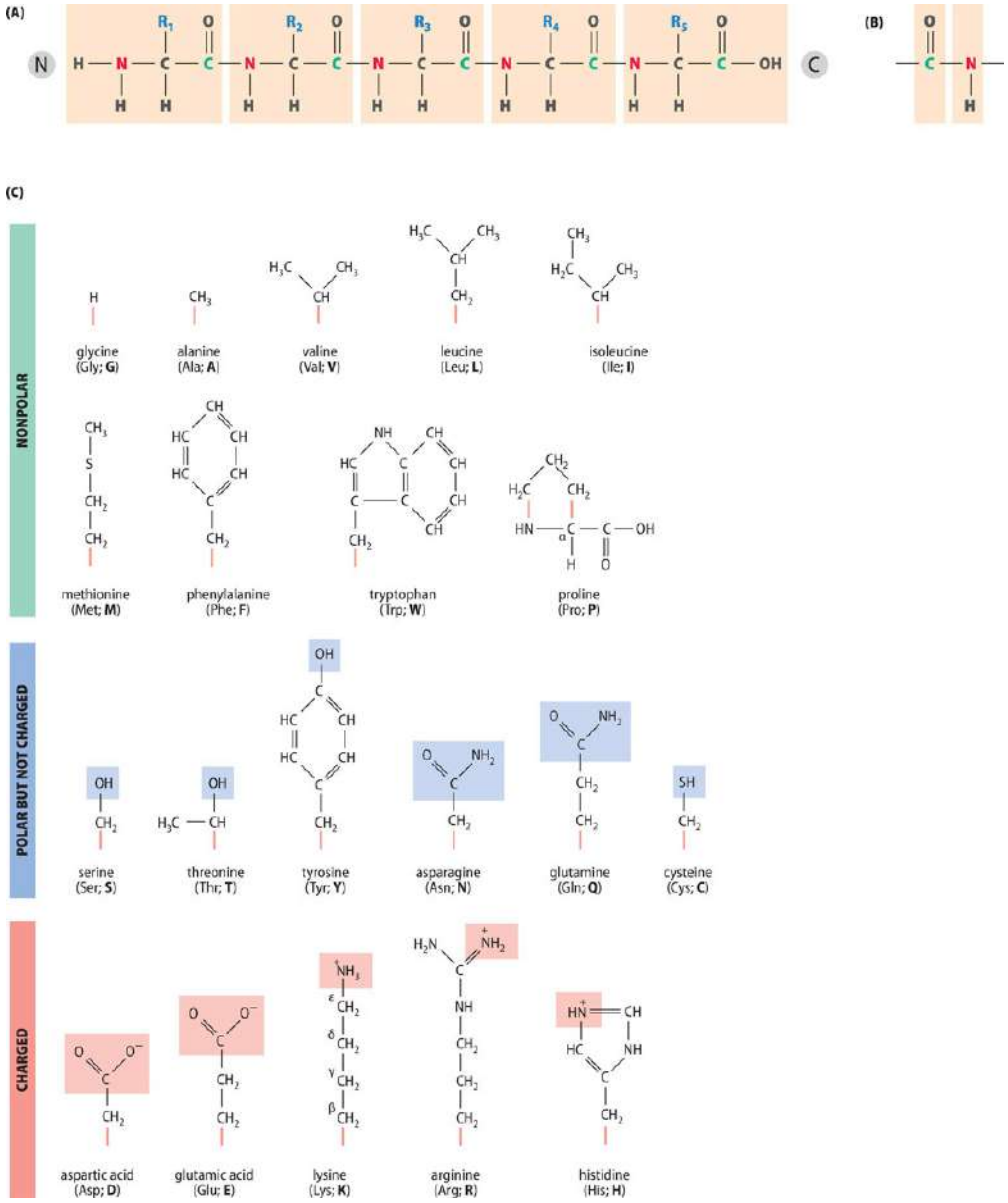


Figure 2.2 Polypeptide and amino acid structure. (A) Polypeptide primary structure. A pentapeptide is shown with its five amino acids highlighted. Here, the left end is called the N-terminal end because the amino acid has a free amino (NH_2) group; the right end is the C-terminal end because the last amino acid has a free carboxyl (COOH) group. The side chains (R_1 to R_5) are variable and determine the identity of the amino acid. They are joined to the central carbon atom of the repeating framework sequence: –NH–CH–CO–. Note that at physiological pH the free amino and carboxyl groups will be charged: NH_3^+ and COO^- respectively. (B) Neighboring amino acids are joined by a peptide bond. A peptide bond is formed by a condensation reaction between the end carboxyl group of one amino acid and the end amino group of another: $-COOH + NH_2- \rightarrow -CONH- + H_2O$. (C) Side chains of the 20 principal amino acids. Red lines represent the covalent bond attaching the side chain to the framework protein structure. Note that the structure of proline is unusual and its full structure is given here because its side chain connects to the nitrogen atom of the framework amino group as well as to the alpha carbon, thereby forming a five-membered ring.

To make a polypeptide, the coding sequence within an mRNA is translated in groups of three nucleotides at a time, called **codons**. There are 64 possible codons (four possible bases at each of three nucleotide positions makes 4 × 4 × 4 permutations). Of these, 61 are used to specify an amino acid; three others signal an end to protein synthesis. The universal **genetic code**, the set of rules that dictate how codons are interpreted, therefore has some redundancy built into it. For example, the amino acid serine can be specified by any of six codons (UCA, UCC, UCG, UCU, AGU, and AGG), and, on average, an amino acid is specified by any of three codons. As a result, nucleotide substitutions within coding DNA quite often do not cause a change of amino acid. We discuss the genetic code in some detail in Section 7.2 when we consider the effects of single nucleotide substitutions.

The process of translation

Translation begins when ribosomes bind to the 5¢ end of an mRNA and then move along the RNA to find a translational start site, the initiation codon—an AUG trinucleotide embedded within the broader, less well defined Kozak consensus sequence (GCC**Pu**CC**AUGG**; the most conserved bases are shown in bold, and Pu represents purine).

The initiation codon is the start of an **open reading frame** of codons that specify successive amino acids in the polypeptide chain (see Box 2.1 for the concept of translational reading frames). As described below, a family of transfer RNAs (tRNAs) is responsible for transporting the correct amino acids to be inserted in the required position of the growing polypeptide chain. Individual types of tRNA carry a specific amino acid; they can recognize and bind to a specific codon, and when they do so they unload their amino acid cargo.

BOX 2.1 TRANSLATIONAL READING FRAMES AND SPLITTING OF CODING SEQUENCES BY INTRONS

TRANSLATIONAL READING FRAMES

In the examples of different translational **reading frames** below, we use sequences of words containing three letters to represent the triplet nature of the genetic code. We designate the reading frames (RF) as 1, 2, or 3 depending on whether the reading frame starts before the first, second, or third nucleotide in the sequence.

Reading frame 1 (RF1) in **Figure 1** makes sense, but a shift to another reading frame produces nonsense. The same principle generally applies to coding sequences. So, for example, if one or two nucleotides are deleted from a coding sequence or there is an insertion of one or two nucleotides, the effect is to produce a **frame-shift** (a change of reading frame) that will result in nonsense.

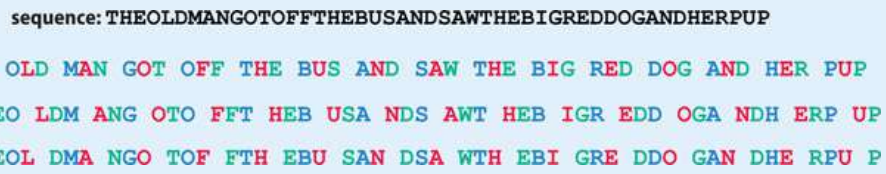


Figure 1 The importance of using the correct translational reading frame. The sequence of letters at the top can be grouped into sets of three (codons) that make sense in reading frame 1 (RF1) but make no sense when using reading

frame 2 (RF2) or reading frame 3 (RF3).

SPLITTING OF CODING SEQUENCES BY INTRONS

At the DNA level, introns may interrupt a coding sequence at one of three types of position: at a point precisely between two codons (a *phase 0 intron*), after the first nucleotide of a codon (a *phase 1 intron*), or after the second nucleotide of a codon (a *phase 2 intron*).

An internal exon may be flanked by introns of the same phase; in an exon like this the number of nucleotides is always exactly divisible by three. Where an exon is flanked by two introns of a different phase, the exon will have a number of nucleotides that is not exactly divisible by three. That can have important consequences when deletions occur within genes (see **Figure 2**).

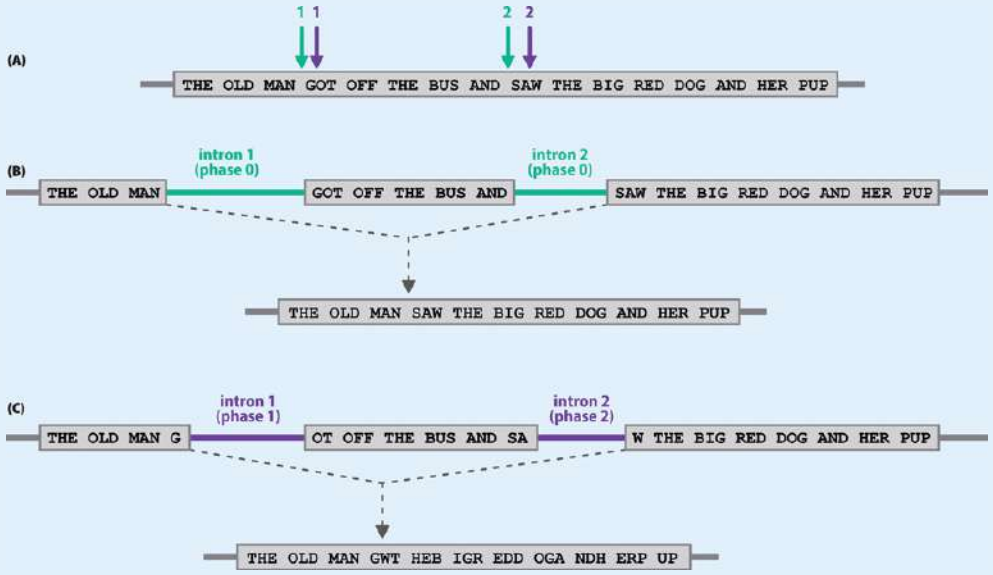


Figure 2 Effects on the translational reading frame caused by the deletion of coding exons. (A) Here we show a coding sequence split at the DNA level by a pair of introns, 1 and 2. Now imagine two alternative possibilities, shown by green and purple arrows. Green arrows indicate two flanking introns of the same phase: in this case both are phase 0 introns, having inserted at analogous positions *between* codons. Purple arrows indicate an alternative where the introns are of different phases, respectively phase 1 for intron 1 and phase 2 for intron 2. B) The green introns result in the central exon having a number of nucleotides that is exactly divisible by three; it can be deleted without an effect on the downstream reading frame. If the exon does not encode a critical component of the protein, the functional consequences may not be too grave. (C) If instead introns 1 and 2 are located as shown in purple, the central exon has a number of nucleotides that cannot be divided exactly by three. If it were to be deleted, the downstream reading frame would be scrambled with a high chance of a premature termination codon, frequently resulting in lack of function.

As each new amino acid is unloaded it is bonded to the previous amino acid so that a polypeptide chain is formed (**Figure 2.3**). The first amino acid has a free NH2 (amino) group and marks the N-terminal end (N) of the polypeptide. The polypeptide chain terminates after the ribosome encounters a **stop codon** (which signifies that the ribosome should disengage from the mRNA, releasing the polypeptide; for translation on cytoplasmic ribosomes, there are three choices of stop codon: UAA,

UAG, or UGA). The last amino acid that was incorporated in the polypeptide chain has a free COOH (carboxyl group) and marks the C-terminal end (C) of the polypeptide.

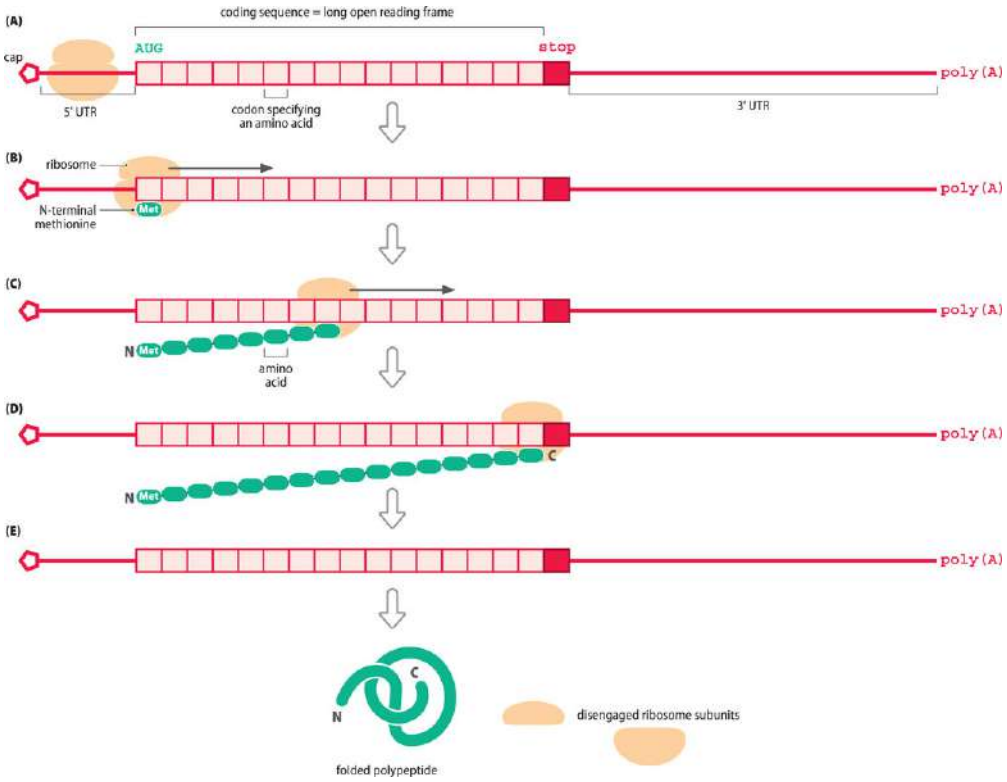


Figure 2.3 The basics of translation. (A) A ribosome attaches to the 5¢ untranslated region (5¢ UTR) of the mRNA and then slides along until (B) it encounters the initiation codon AUG, at which point a methionine-bearing transfer RNA (not shown) engages with the AUG codon and deposits its methionine cargo (green bar, labeled MET). (C) The ribosome continues to move along the mRNA and as it encounters each codon in turn a specific amino-acid-bearing tRNA is recruited to recognize the codon and to deposit its amino acid, according to the *genetic code*. The ribosome catalyses the formation of a *peptide bond* (Figure 2.2B) between each new amino acid and the last amino acid, forming a polypeptide chain (shown here, for convenience, as a series of joined green ovals). (D) Finally, the ribosome encounters a stop codon, at which point (E) the ribosome falls off the mRNA and dissociates into its two subunits, releasing the completed polypeptide. The polypeptide undergoes *post-translational modification* as described in the text, which may sometimes involve cleavage at the N-terminal end so that methionine may not be the N-terminal amino acid in the mature polypeptide.

Transfer RNA as an adaptor RNA

Transfer RNAs have a classic cloverleaf structure resulting from intramolecular hydrogen bonding (**Figure 2.4A**). They serve as adaptor RNAs because their job is to base pair with mRNAs and help decode the coding sequence messages carried by mRNAs. The base pairing is confined to a three-nucleotide sequence in the tRNA called an *anticodon*, which is complementary in sequence to a codon. According to the identity of their anticodons, different tRNAs carry different amino acids covalently linked to their 3¢ ends. Through base pairing between codon and anticodon, individual

amino acids can be sequentially ordered according to the sequence of codons in an mRNA, and sequentially linked together to form a polypeptide chain (see Figure 2.4B).

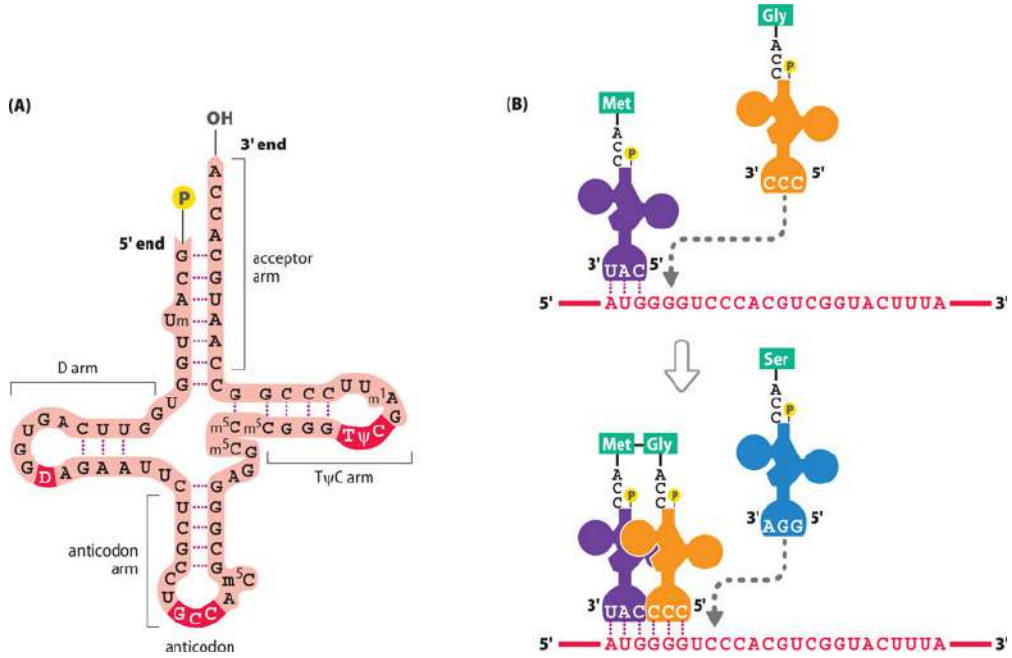


Figure 2.4 Transfer RNA structure, and its role as an adaptor RNA in translation. (A) Transfer RNA structure. The tRNA^{Gly} shown here illustrates the classical cloverleaf tRNA structure. Intramolecular base pairing produces three arms terminating in a loop plus an acceptor arm formed by pairing of 5¢ and 3¢ end sequences. The latter ends in a -CAA trinucleotide and covalently binds a specific amino acid. The three nucleotides at the centre of the middle loop form the *anticodon,* which identifies the tRNA according to the amino acid it will bear. Minor nucleotides are: D, 5,6-dihydrouridine; Y, pseudouridine (5-ribosyluracil); m5C, 5-methylcytidine; m1A, 1-methyladenosine; Um, 2¢-*O*-methyluridine. (B) Role of adaptor RNA. Different tRNAs carry different amino acids, according to the type of anticodon they bear. As a ribosome traverses an mRNA it identifies the AUG initation codon. A methionine-bearing tRNA with the complementary CAU anticodon sequences then engages with the ribosome so that the CAU anticodon base pairs with the AUG codon. Note: for ease of illustration we show the tRNAs in the opposite orientation to the standard form shown in (A), with the acceptor arm on the left, not the right. Thereafter, a tRNA bearing glycine engages the second codon, GGG, by base pairing with its CCC anticodon. The ribosome's peptidyltransferase then forms a peptide bond between the N-terminal methionine and glycine. The ribosome moves along by one codon and the tRNA^{Met} is cleaved so that it can be reused and the process continues with an incoming tRNA carrying a serine and an anticodon GGA to bind to the third codon UCC, after which the incoming serine will be covalently bonded to the glycine by the ribosome's peptidyltransferase.

Untranslated regions and 5′ cap and 3′ poly(A) termini

As illustrated in Figure 2.3, each mature mRNA has a large central coding DNA sequence flanked by two **untranslated regions**, a short 5¢ untranslated region (5¢ UTR) and a rather longer 3¢ untranslated region (3¢ UTR). The untranslated regions regulate mRNA stability and contain regulatory sequences that are important in determining how genes are expressed.

As well as sequences copied from the gene sequence, mRNA molecules usually also have end sequences added post-transcriptionally to the pre-mRNA. At the 5¢ end a specialized cap is added: 7-methylguanosine linked to the first nucleotide by a distinctive 5¢-5¢ phosphodiester bond (instead of a normal 5¢-3¢ phosphodiester bond). The cap protects the transcripts against 5¢ ® 3¢ exonuclease attack and facilitates transport to the cytoplasm and ribosome attachment. At the 3¢ end a dedicated poly(A) polymerase sequentially adds adenylate (AMP) residues to give a poly(A) tail, about 150–200 nucleotides long. The poly(A) helps in transporting mRNA to the cytoplasm, facilitates binding to ribosomes, and is also important in stabilizing mRNAs.

From newly synthesized polypeptide to mature protein

The journey from newly synthesized polypeptide released from the ribosome to fully mature protein requires several steps. The polypeptide typically undergoes post-translational cleavage and chemical modification. Polypeptides also need to fold properly, and they often bind to other polypeptides as part of a multisub-unit protein. And then there is a need to be transported to the correct intracellular or extracellular location.

Chemical modification

We describe below one type of chemical modification that involves cross-linking between two cysteine residues within the same polypeptide or on different polypeptides. Often, however, chemical modification involves the simple covalent addition of chemical groups to polypeptides or proteins. Sometimes small chemical groups are attached to the side chains of specific amino acids (**Table 2.2**). These groups can sometimes be particularly important in the structure of a protein (as in the case of collagens, which have high levels of hydroxyproline and hydroxylysine).

TABLE 2.2 COMMON TYPES OF CHEMICAL MODIFICATION OF PROTEINS BY COVALENT ADDITION OF CHEMICAL GROUPS TO A SIDE CHAIN

Type of chemical modification	Target amino acids	Comments
ADDITION OF SMALL CHEMICAL GROUP		
Hydroxylation	Pro; Lys; Asp	can play important structural roles
Carboxylation	Glu	especially in some blood clotting factors
Methylation	Lys	specialized enzymes can add or remove the methyl, acetyl, or phosphate group, causing the protein to switch between two states, with functional consequences
Acetylation	Lys	
Phosphorylation	Tyr;Ser;Thr	
ADDITION OF COMPLEX CARBOHYDRATE OF LIPID GROUP		
N-glycosylation	Asn	added to the amino group of Asn in endoplasmic reticulum and Golgi apparatus
O-glycosylation	Ser;Thr; Hydroxylysine	added to the side-chain hydroxyl group; takes place in Golgi apparatus

Type of chemical modification	Target amino acids	Comments
N-lipidation	Gly	added to the amino group of an N-terminal glycine; promotes protein-membrane interactions
S-lipidation	Cys	a palmitoyl or prenyl group is added to the thiol of the cysteine. Often helps anchor proteins in a membrane

In other cases, dedicated enzymes add or remove small chemical groups to act as switches that convert a protein from one functional state to another. Thus, specific kinases can add a phosphate group that can be subsequently removed by a dedicated phosphatase. The change between phosphorylated and dephosphorylated states can result in a major conformational change that affects how the protein functions. Similarly, methyltransferases and acetyltransferases add methyl or acetyl groups that can be removed by the respective demethylases and deacetylases. As we will see in Chapter 6, they are particularly important in modifying histone proteins to change the conformation of chromatin and thereby alter gene expression.

In yet other cases, proteins can be modified by covalently attaching complex carbohydrates or lipids to a polypeptide backbone. Thus, for example, secreted proteins and proteins destined to be part of the excretory process of cells routinely have oligosaccharides attached to the side chains of specific amino acids. Different types of lipids are also often added to membrane proteins (see Table 2.2).

Folding

The amino acid sequence, the primary structure, dictates the pattern of folding, but certain regions of polypeptides adopt types of secondary structure important in protein folding (**Box 2.2** gives an outline of protein structure). Until correct folding has been achieved, a protein is unstable; different chaperone molecules help with the folding process (careful supervision is needed because partly folded or misfolded proteins can be toxic to cells).

BOX 2.2 A BRIEF OUTLINE OF PROTEIN STRUCTURE

Four different levels of structure are recognized:

- primary structure—the linear sequence of amino acids in constituent polypeptides
- secondary structure—the path that a polypeptide backbone follows within local regions of the primary structure
- tertiary structure—the overall three-dimensional structure of a polypeptide
- quaternary structure—the aggregate structure of a multimeric protein (composed of two or more polypeptide subunits that may be of more than one type).

ELEMENTS OF SECONDARY STRUCTURE

Secondary structure is notably shaped by intramolecular hydrogen bonding. The *α-helix*, for example, is a rigid cylinder stabilized by hydrogen bonding between the carbonyl oxygen of a

peptide bond and the hydrogen atom of the amino nitrogen of a peptide bond located four amino acids away (Figure 1A). a-Helices often occur in transcription factors and other proteins that perform key cellular functions.

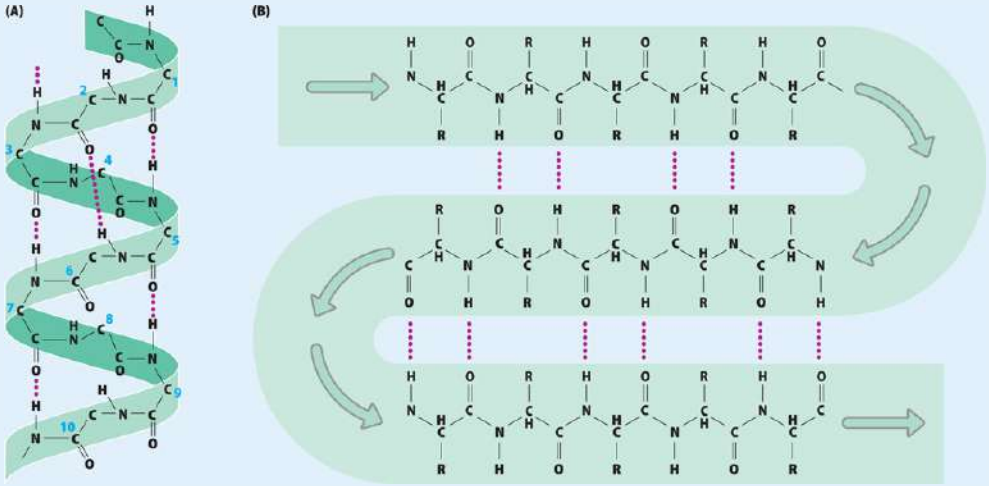


Figure 1 Elements of protein secondary structure. (A) An a-helix. The solid rod structure is stabilized by hydrogen bonding between the oxygen of the carbonyl group (C = O) of each peptide bond and the hydrogen on the peptide bond amide group (NH) of the fourth amino acid away, yielding 3.6 amino acids per turn of the helix. The side chains of each amino acid are located on the outside of the helix; there is almost no free space within the helix. Note: only the backbone of the polypeptide is shown, and some bonds have been omitted for clarity. (B) A b-sheet (also called a b-pleated sheet). Here, hydrogen bonding occurs between the carbonyl oxygens and amide hydrogens on adjacent segments of a sheet that may be composed either of parallel segments of the polypeptide chain or, as shown here, of antiparallel segments (arrows mark the direction of travel from N-terminus to C-terminus) ……., hydrogen bond.

In the *β-sheet* (also called the *β-pleated sheet*), the hydrogen bonds occur between opposed peptide bonds in parallel or antiparallel segments of the same polypeptide chain (Figure 1B). b-Sheets occur—often together with a-helices—at the core of most globular proteins.

The *β-turn* involves hydrogen bonding between the peptide-bond carbonyl (C = O) of one amino acid and the peptide-bond NH group of an amino acid located only three places farther along. The resulting hairpin turn allows an abrupt change in the direction of a polypeptide, enabling compact globular shapes to be achieved. b-Turns can connect neighboring segments in a b-sheet, when the polypeptide strand has to undergo a sharp turn.

When placed in an aqueous environment, proteins are stabilized by having amino acids with hydrophobic side chains located in the interior of the protein, whereas hydrophilic amino acids tend to be located toward the surface. For many proteins, notably globular proteins, the folding pattern is also stabilized by a form of covalent cross-linking that can occur between certain distantly located cysteine residues—the sulfhydryl groups of the cysteine side chains interact to form a disulfide bond (alternatively called a disulfide bridge—see Figure 2.5).

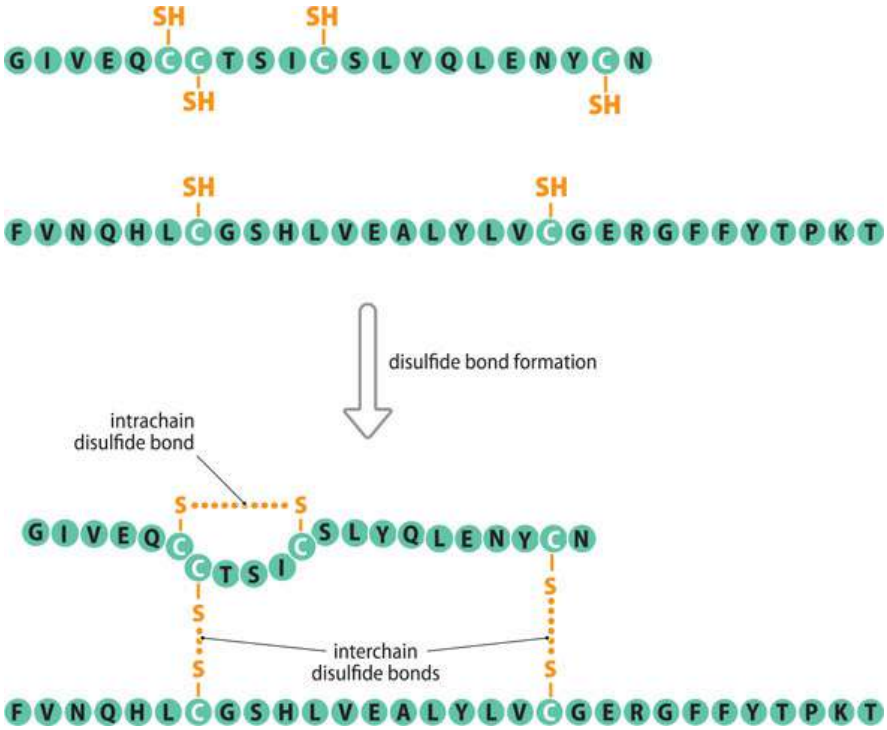


Figure 2.5 Intrachain and interchain disulfide bridges in human insulin. Human insulin is composed of two peptide chains, an A chain with 21 amino acids, and a B chain with 30 amino acids. Disulfide bridges (−S−S−) form by a condensation reaction between the sulfhydryl (−SH) groups on the side chains of cysteine residues. They form between the side chains of cysteines at positions 6 and 11 within the insulin A chain, and also between cysteine side chains on the insulin A and B chains. Note that here all the cysteines participate in disulfide bonding, which is unusual. When disulfide bonding occurs in large proteins, only certain cysteine residues are involved.

Cleavage and transport

The initial polypeptide normally undergoes some type of N-terminal cleavage. Sometimes just the N-terminal methionine is removed. But for proteins secreted from cells, the polypeptide precursor carries an N-terminal leader sequence (signal peptide) that is required to assist the protein to cross the plasma membrane, after which the signal peptide is cleaved at the membrane, releasing the mature protein. (The signal peptide, often 10–30 amino acids in length, carries multiple hydrophobic amino acids.)

Other short internal peptide sequences can act simply as address labels for transporting proteins to the nucleus, mitochondria, plasma membrane, and so on. They are retained in the mature protein.

Binding of multiple polypeptide chains

Proteins are often made of two or more polypeptide subunits. Occasionally, constituent polypeptides are covalently linked with disulfide bridges (as in the case of joining the different chains of immunoglobulins; see Figure 4.10 on page 99). Often, however, constituent polypeptides are held together mainly by noncova-lent bonds, including nonpolar interactions and hydrogen bonds. For example, hemoglobins are tetramers, composed of two copies each of two different globin chains that

associate in this way. Collagens provide a good example of very intimate structural association between polypeptides, consisting of three chains (two of one type; one of another) wrapped round each other to form a triple helix.

2.2 RNA GENES AND NONCODING RNA

The majority of our genes are **RNA genes**, genes devoted to making functional **noncoding RNA** (**ncRNA**) as their end product. (The latest GENCODE data – RELEASE 40, April 2022 – revealed a total of 26 372 human RNA genes and 19 988 protein-coding genes). The vast majority of the RNA genes regulate gene expression in some way, or directly assist in the expression of protein-coding genes, and proteins remain the main functional endpoint in cells.

Like proteins (and mRNA), noncoding RNAs are made as precursors that often undergo enzymatic cleavage to become mature gene expression products. They are also subject to chemical modification: minority bases such as dihydrouri-dine or pseudouridine and various methylated bases are quite common—see Figure 2.4A for some examples in a tRNA.

Until quite recently, ncRNAs were largely viewed as having important but rather dull functions. For the most part, they seemed to act as ubiquitous accessory molecules that worked directly or indirectly in protein production. After ribosomal and transfer RNAs, we came to know about various other ubiquitous ncRNAs that mostly work in RNA maturation: spliceosomal small nuclear RNAs (snRNAs); small nucleolar RNAs (snoRNAs) that chemically modify specific bases in rRNA; small Cajal-body RNAs (scaRNAs) that chemically modify spliceosomal snRNA; and certain RNA enzymes (ribozymes) that cleave tRNA and rRNA precursors. All of these types of RNA can be viewed as accessory molecules needed, like rRNA and tRNA, to support protein synthesis in general. In stark contrast to RNA, proteins were viewed as the functionally important endpoints of genetic information, the exciting pacesetters that performed myriad roles in cells.

The view that noncoding RNAs (ncRNAs) are mostly ubiquitous accessory molecules that assist general protein synthesis is no longer tenable. Over the past two decades we have become progressively more aware of the functional diversity of ncRNA and of the many thousands of ncRNA genes in our genome. Multiple new classes of regulatory RNAs have very recently been discovered to be expressed in certain cell types only, or at certain stages of development. Working out what they do has become an exciting area of research.

With hindsight, perhaps we should not be so surprised at the functional diversity of RNA. DNA is simply a self-replicating repository of genetic information, but RNA can serve this function (in the case of RNA viruses) and can also have catalytic functions. In the “RNA world” hypothesis RNA is viewed as the original genetic material and as also being capable of executive functions before DNA and proteins developed. That is possible because, unlike naked double-stranded DNA (which has a comparatively rigid structure), single-stranded RNA has a very flexible structure and can form complex shapes by intramolcular hydrogen bonding, as described below. As will be described in later chapters, the relatively recent understanding of just what RNA does in cells and how it can be manipulated is driving some important advances in medicine. Mutations in certain RNA genes are now known to underlie some genetic disorders and cancers, and RNA therapeutics offers important new approaches to treating disease.

The extraordinary secondary structure and versatility of RNA

The primary structure of nucleic acids and proteins is the sequence of nucleotides or amino acids that defines their identity; however, higher levels of structure determine how they work in cells. Single-stranded RNA molecules are much more flexible than naked double-stranded DNA, and like proteins they have a very high degree of secondary structure where intramolecular hydrogen bonding causes local alterations in structure.

The secondary structure of single-stranded RNA depends on base pairing between complementary sequences on the same RNA strand. Intervening sequences that do not engage in base pairing will loop out, producing stem-loop structures (called hairpins when the loop is short)—see **Figure 2.6**. Higher-level structures can form when, for example, a sequence within the stem of one loop base pairs with another sequence, and extraordinarily intricate structures can develop. Note that base pairing in RNA includes G–U base pairs as well as more stable A–U and G–C base pairs.

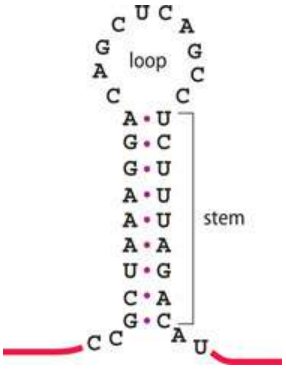


Figure 2.6 A stem-loop structure. This is formed when the RNA folds back on itself so that two short regions can base pair to form the stem while a small intervening sequence loops out. Note that G–U base pairs form in RNA, in addition to G–C and A–U base pairs. Related structures but with shorter stems are important in tRNA structure as shown in Figure 2.4A.

Stem-loop structures in RNA have different functions. As described in Chapter 6, they can serve as recognition elements for binding regulatory proteins, and they are crucially important in determining the overall structure of an RNA that can be important for function.

In general, because of the flexible structure of single-stranded RNA, different RNAs can adopt different shapes according to the base sequence; this enables them to do different jobs, such as working as enzymes. Many different classes of RNA enzyme (ribozyme) are known in nature, and some originated very early in evolution. For example, the catalytic activity of the ribosome (the peptidyltransferase responsible for adding amino acids to the growing poly-peptide chain) is due solely to the large RNA (28S rRNA) present in the large subunit. In recent years RNAs have been found to work in a large variety of roles (**Figure 2.7**).

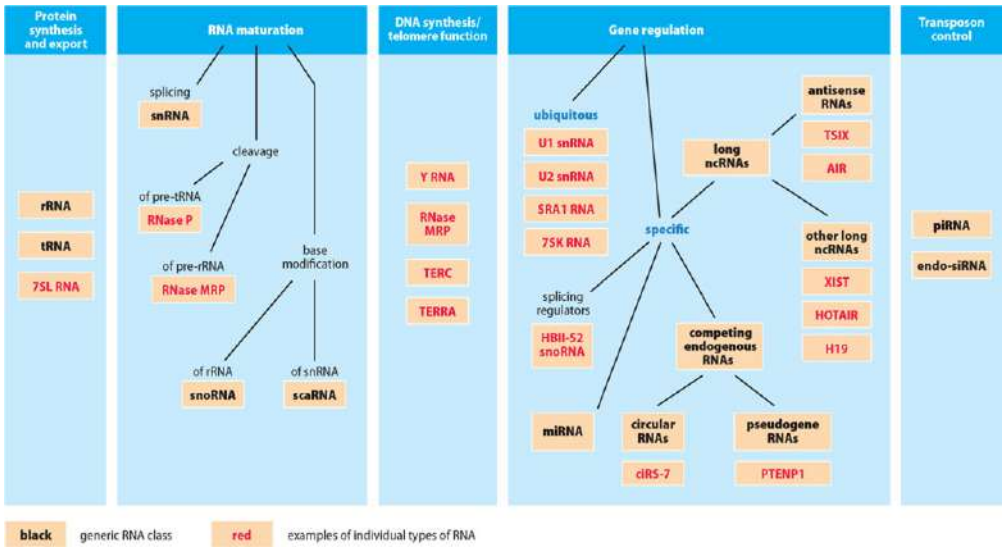


Figure 2.7 The versatility of noncoding RNA. The two panels on the left show ubiquitously expressed RNAs that are important in generally assisting protein production and export, including RNA families that supervise the maturation of other RNAs, notably: small nuclear RNA (snRNA); small nucleolar RNA (snoRNA); and small Cajal-body RNA (scaRNA). The central panel includes RNAs involved in DNA replication (the ribozyme RNase MRP has a crucial role in initiating mtDNA replication, as well as in cleaving pre-rRNA), and developmentally regulated telomere regulators (TERC is the RNA component of telomerase; TERRA is telomere RNA). Diverse classes of noncoding RNA regulate gene expression. In addition to the listed ubiquitous RNAs that have general roles in transcription, many classes of RNA regulate *specific* target genes and are typically restricted in expression. They work at different levels: transcription (such as antisense RNAs), splicing, and translation (notably miRNAs, which bind to certain regulatory sequences in the untranslated regions of target mRNAs). Some RNAs, notably the highly prevalent class of circular RNAs, regulate the interaction between miRNAs and their targets. piRNAs, and to a smaller extent endogenous short interfering RNAs (endo-siRNA), are responsible for silencing transposable elements in germline cells. We describe how RNAs regulate gene expression in detail in Chapter 6.

RNAs that act as specific regulators: from quirky exceptions to the mainstream

The first examples of more specific regulatory RNAs were discovered more than 20 years ago. For a long time they were considered interesting but exceptional cases. They included RNAs working in epigenetic regulation to produce monoallelic gene expression. For most genes both the maternal and paternal allele are normally expressed, but for a few genes it is *normal* that only one of the two parental alleles is expressed. Some of our genes are *imprinted* so that, according to the specific gene, either the maternal allele or the paternal is consistently expressed, while the other allele is silenced. And in women (and female mammals), genes on one of the two X chromosomes, either the maternal or the paternal X chromosome chosen at random, are *normally* silenced (X-chromosome inactivation). We describe the underlying mechanisms in Chapter 6.

We now know that there are many thousands of different RNA genes in our genome. Many of these genes make regulatory ncRNAs that are expressed in certain cell types only, including some large families of long noncoding RNAs and tiny noncoding RNAs.

Long noncoding RNAs

In addition to the very few ribosomal RNAs, there are a very large number of long noncoding RNAs, the great majority of which are associated with chromatin and act as regulators of gene expression. They come in two broad classes. **Antisense RNAs** are transcribed using the *sense* strand of a gene as a template and are not subject to cleavage and RNA splicing. As a result they can be quite large, often many thousands of nucleotides long. They work by binding to the complementary sense RNA produced from the gene, downregulating gene expression.

A second class of long regulatory RNAs are formed from primary transcripts that are typically processed like the primary transcripts of protein-coding genes (and so normally undergo RNA splicing). Many of these RNAs regulate neighboring genes, but some control the expression of genes on other chromosomes. We consider the details of how they work in Chapter 6.

Tiny noncoding RNAs

Thousands of tiny noncoding RNAs (less than 35 nucleotides long) also work in human cells. They include many microRNAs (miRNAs) that are usually 20–22 nucleotides long and are expressed in defined cell types or at specific stages of early development. As described in Chapter 6, a miRNA works by recognizing and binding to defined target regulatory sequences present in specific mRNAs in order to downregulate their expression. MicroRNAs are important in a wide variety of different cellular processes.

Human germ cells also make many thousands of different 26–32-nucleotide Piwi protein-interacting RNAs (piRNAs). The piRNAs work in germ cells to damp down excess activity of **transposons** (mobile DNA elements). Active mobile elements in the human genome can make a copy that migrates to a new location in our genome and can be harmful (by disrupting genes or inappropriately activating some types of cancer gene).

2.3 WORKING OUT THE DETAILS OF OUR GENOME AND WHAT THEY MEAN

The human genome consists of 25 different DNA molecules partitioned between two physically separate genomes, one in the nucleus and one in the mitochondria. In the nucleus there are either 23 or 24 different types of linear DNA molecule (one each for the *different* types of chromosome: 23 in female cells or 24 in male cells). The chromosomal DNA molecules are immensely long (ranging in size from 48 Mb to 249 Mb). In the mitochondria there is just one type of DNA molecule: a comparatively tiny circular DNA just 16.6 kilobases (kb) long, roughly 1/10 000 of the size of an average nuclear DNA molecule. Unlike the chromosomal DNA molecules (each present in only two copies in diploid cells), there are many mitochondrial DNA copies in a cell, and the copy number can vary very significantly according to the type of cell.

In what was a heroic effort at the time, the mitochondrial DNA (often called the mitochondrial genome) was sequenced by a single research team in Cambridge, UK, as far back as 1981. Despite its small size, it is packed with genes. The complexity of the nuclear genome—roughly 200 000 times the

size of the mitochondrial genome—posed a much more difficult challenge. That would require an international collaboration between many research teams, as described below.

Working out the nucleotide sequence was only the first step. The next challenge, which is still continuing and may take decades, is to work out the details of how our genome functions and what all the component sequences do.

The Human Genome Project: working out the details of the nuclear genome

For decades, the only available map of the nuclear genome was a low-resolution physical map based on chromosome banding. Chromosomes can be stained with certain dyes, such as Giemsa, to reveal an alternating pattern of dark and light bands for each chromosome, as represented by the image shown in **Figure 2.8**.

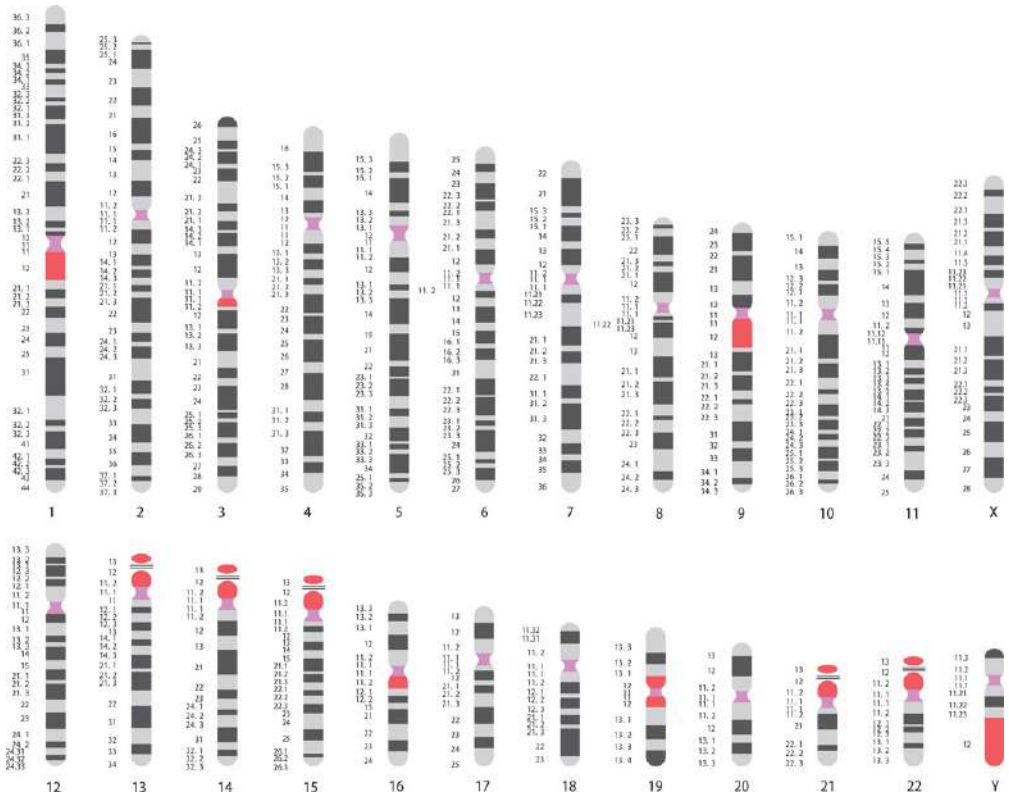


Figure 2.8 Ideogram showing a 550-band Giemsa banding pattern and constitutive heterochromatin within human metaphase chromosomes. Dark bands represent DNA regions where there is a low density of G–C base pairs and a generally low density of exons and genes. Pale bands represent DNA regions where there is a high density of G–C base pairs and a generally high density of exons and genes. Centromeric heterochromatin is illustrated by mauve blocks; non-centromeric and non-telomeric constitutive heterochromatin is shown as bright red blocks. Note the large amounts of non-centromeric heterochromatin on the Y chromosome, the short arms of the acrocentric chromosomes (13, 14, 15, 21, and 22), and on chromosomes 1, 9, and 16. Numbers to the left are the numbers of individual chromosome bands; for the nomenclature of chromosome banding, see Box 7.2 on pages 204–5.

We describe the methodology and terminology of human chromosome banding in Box 7.2. For now, there are two salient points to note. First, the alternating pattern of bands reflects different staining intensities. That in turn reflects differences in chromatin organization along chromosomes (as

a result of differences in base composition), and differences in gene and exon density (see the legend to Figure 2.8). Secondly, the resolution of the map is low—in even a high-resolution chromosome map the average size of a band is several megabases of DNA. What was needed was a map with a 1 bp resolution, a DNA sequence map.

The principal objective of the international Human Genome Project (HGP) was to obtain a *reference sequence* for our nuclear genome, that is, the aggregated DNA sequences of each of the 24 different human chromosomes. It is necessarily a reference sequence: if we assume 8 billion people on the planet, and without even considering somatic variation, there are at least 16 billion different human genomes (each of us has inherited two different genomes: a maternal genome and a paternal genome that also show differences from the DNA of our parents because of meiosis). The HGP recruited a small number of individuals who donated blood cells to provide genomic DNA for the project. Ultimately, for each chromosome, a map was constructed based on many DNA clones with long inserts that could be ordered as a series of DNA clones with partly overlapping DNA sequences. Finally, the DNAs from selected clones were sequenced and used to build chromosome-wide DNA sequences.

Not all regions of DNA were sequenced: a low priority was given to long regions composed largely of highly repetitive DNA repeats that were technically difficult to sequence. They notably included regions of **heterochromatin**, the parts of the genome where the chromatin of cells is highly condensed throughout the cell cycle—see Box 2.3 for an overview of heterochromatin.

Instead, the focus was on sequencing the gene-rich **euchromatin** component of our genome. A draft sequence for the human nuclear euchromatin genome was published in 2001 (after collation of all the different chromosome DNA sequences). Thereafter, almost complete sequences of the euchromatin region of all 24 nuclear DNA molecules were obtained and published by 2003–2004. The DNA of the euchromatin component was found to represent about 93 % of the nuclear genome. Subsequently, significant components of heterochromatin DNA have been sequenced.

BOX 2.3 AN OVERVIEW OF EUCHROMATIN AND HETEROCHROMATIN

Like other complex genomes, the human nuclear genome is composed of regions of gene-poor **heterochromatin** (where the DNA has a very highly condensed structure that acts as a barrier to transcription factors), and gene-rich **euchromatin** (where the DNA is more open and generally more accessible to transcription factors).

There is some variability in both cases. For euchromatin, the transcriptional activity of euchromatin can vary between cells, notably between cells of different types. In addition to genes expressed in essentially all nucleated cells, many genes in euchromatin show restricted expression; to allow tissue-specific expression, specific regions of euchromatin are induced to be more condensed and transcriptionally inactive in some cells but in other cells the equivalent gene has an open structure accessible to transcription factors.

For heterochromatin, the variability depends on the extent to which the highly condensed structure is consistently maintained or is prone to being altered by an epigenetic mechanism to allow transcription under certain circumstances. As a result, two types of heterochromatin have been distinguished, as listed below.

- ***Facultative heterochromatin.*** In this case, the same specific regions of DNA can be very highly condensed in some circumstances but be induced to undergo a change in structure to behave as euchromatin in other circumstances. For example, because of an epigenetic phenomenon known as X-inactivation, the two X chromosomes in a woman are very different: one where most of the X is highly condensed and transcriptionally inactive, and the other where almost all of the X has a relaxed euchromatin structure.
- ***Constitutive heterochromatin.*** This type of gene-poor chromatin is consistently condensed, at least in somatic cells, and accounts for ~7 % of the nuclear genome. It is found at the centromeres and telomeres, and also over significant other regions of certain autosomes (notably human chromosomes 1, 9, 13–16, 19, 21, 22) and the Y chromosome (see Figure 2.8). Note: constitutive heterochromatin does contain some genes that are expressed in germline cells. An example is the *DUX4* gene located in telomeric heterochromatin at 4q35.2. If inappropriately expressed in muscle cells, it can cause facioscapulohumeral muscular dystrophy, as described in Section 6.3.

What the sequence didn't tell us and the goal of identifying all functional human DNA sequences

With hindsight, we can appreciate that obtaining the human genome sequence was the easy part. The hard part is to work out what the sequence means. We now know that much of our genome is composed of repetitive sequences. Many genes, for example, are members of families of closely related genes that also often contain *pseudogenes,* inactive copies of functional genes. And a good deal of the genome is composed of highly repetitive noncoding DNA sequences. Because much of the genome sequence does not appear to be critically important, a priority in the "post-genome era" has been to identify all the functionally important DNA sequences and understand how they work.

Protein-coding genes

When our genome sequence was first obtained, it was widely anticipated that new (previously unstudied) genes would be revealed. And indeed, many new human protein-coding genes were soon discovered using computer-based analyses (coding DNA sequences are usually easy to identify because they have long open reading frames and are highly conserved during evolution). Various programs can scan the genome for longer-than-expected open reading frames (as expected of coding DNAs). BLAST programs (described below) can then compare candidate human coding DNAs to seek related sequences in the genomes of other mammals (such as mouse). BLAST programs that rely on translating the putative coding DNA in all six reading frames (three each for the two DNA strands) and then comparing them with similarly translated mammalian genomes, were especially effective (protein sequences are even more highly conserved in evolution than coding DNA—see **Figure 2.9** for the example of human and mouse p53 protein sequences).

score	expect	method	identities	positives	gaps
574 bits(1480)	0.0	Compositional matrix adjust.	304/393(77%)	326/393(82%)	6/393(1%)

```
Query  1    MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP  60
            MEE QSD S+E PLSQETFS LWKLLP  ++L P P   MDDL+L P D+E++F    GP
Sbjct  4    MEESQSDISLELPLSQETFSGLWKLLPPEDIL-PSP-HCMDDLLL-PQDVEEFFE---GP  57

Query  61   DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK  120
             EA R+  A     P    P P APAPA  WPLSS VPSQKTYQG+YGF LGFL SGTAK
Sbjct  58   SEALRVSGAPAAQDPVTETPGPVAPAPATPWPLSSFVPSQKTYQGNYGFHLGFLQSGTAK  117

Query  121  SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE  180
            SV CTYSP LNK+FCQLAKTCPVQLWV +TPP G+RVRAMAIYK+SQHMTEVVRRCPHHE
Sbjct  118  SVMCTYSPPLNKLFCQLAKTCPVQLWVSATPPAGSRVRAMAIYKKSQHMTEVVRRCPHHE  177

Query  181  RCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNYMCNS  240
            RCSD DGLAPPQHLIRVEGNL  EYL+DR TFRHSVVVPYEPPE GS+ TTIHY YMCNS
Sbjct  178  RCSDGDGLAPPQHLIRVEGNLYPEYLEDRQTFRHSVVVPYEPPEAGSEYTTIHYKYMCNS  237

Query  241  SCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRDRRTEEENLRKKGEPHHELP  300
            SCMGGMNRRPILTIITLEDSSGNLLGR+SFEVRVCACPGRDRRTEEEN RKK     ELP
Sbjct  238  SCMGGMNRRPILTIITLEDSSGNLLGRDSFEVRVCACPGRDRRTEEENFRKKEVLCPELP  297

Query  301  PGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG  360
            PGS KRALP  TS+SP  KKKPLDGEYFTL+IRGR+RFEMFRELNEALELKDA A +E G
Sbjct  298  PGSAKRALPTCTSASPPQKKKPLDGEYFTLKIRGRKRFEMFRELNEALELKDAHATEESG  357

Query  361  GSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD  393
             SRAHSS+LK+KKGQSTSRHKK M K  GPDSD
Sbjct  358  DSRAHSSYLKTKKGQSTSRHKKTMVKKVGPDSD  390
```

Figure 2.9 Unlike RNA sequences, protein sequences are often highly conserved in evolution: the example of human and mouse p53 proteins. The alignment was generated by the BLASTP program that compared a query sequence, the human p53 protein, against the mouse p53 sequence, the subject (Sbjct) sequence. The alignment is slightly skewed because the mouse p53 sequence has an extra three amino acids at its N-terminus, but lacks some amino acids found in the human p53 sequence, such as the EDP at positions 56–58 in human p53). The middle lines between the query and subject lines show identical residues (at 302 out of the 393 matched positions). The + symbol in the middle lines denote functionally similar amino acids; they occur here at 22 positions, which when added to 304 identities gives a total of 326 positive matches out of 393.

RNA genes and regulatory elements

The vast majority of known human RNA genes and regulatory sequences (such as promoters, enhancers, and so on) were initially not identified from the genome sequence. RNA genes lack open reading frames, some of them are tiny sequences, and unlike the polypeptide sequences of proteins, RNA sequences are often poorly conserved during evolution (in RNA molecules the shape is important, but the sequence less so); but nucleotides are important in maintaining the RNA shape or in ensuring correct binding to interacting molecules and so are often well-conserved. Regulatory elements are clusters of tiny sequences that are often not as well-conserved as coding DNA, making them often difficult to identify by computer programs in the past. Our knowledge of human regulatory elements was therefore also limited at the time when the human genome sequence was reported in 2003.

For follow-up projects to hunt down all functional human DNA sequences, the major priority became to identify and catalog all RNA transcripts and also all regulatory sequences (working at either the DNA level, or at the RNA level). The projects used genome-wide transcription analyses,

evolutionary sequence comparisons and functional assays, all underpinned by bioinformatic analyses. The most prominent international study—the ENCODE (**Enc**yclopedia **o**f **D**NA **E**lements) Project reported its findings in 2012, and one of its most important findings was that RNA transcription is pervasive. And, as detailed in Section 2.5, it has become clear that there are significantly more RNA genes in the human genome than protein-coding genes.

2.4 A QUICK TOUR OF SOME ELECTRONIC RESOURCES USED TO INTERROGATE THE HUMAN GENOME SEQUENCE AND GENE PRODUCTS

A wide variety of databases and computer programs currently provide a wealth of information on the human genome, human genes, and gene products (**Table 2.3**). Genome browsers help users navigate a sequenced genome by programs that employ graphical user interfaces to portray genome information for selected chromosomes and subchromosomal regions. The characteristics (genes, exons, transcripts, and so on) of a selected human chromosome or chromosome region can be tracked, moving from large scale to nucleotide scale, with click-over facilities to identify the characteristics and download the sequences of genes and associated exons, RNAs, and proteins. The principal browsers are listed in Table 2.3.

TABLE 2.3 SOME OF THE PRINCIPAL ELECTRONIC RESOURCES FOR INTERROGATING THE HUMAN GENOME, HUMAN GENES AND GENE PRODUCTS

Resource use	Popular resources	Website address
Gateways to multiple electronic resources	Human Genome Resources HGNC portal NIH National Library of Medicine	http://www.ncbi.nlm.nih.gov/genome/guide/human/ http://www.genenames.org/ http://www.ncbi.nlm.nih.gov/gquery/
Reference nucleotide and protein sequences	RefSeq (for mRNAs, ncRNAs and proteins) RefSeqGene (for genes)	http://www.ncbi.nlm.nih.gov/refseq/ http://www.ncbi.nlm.nih.gov/refseq/rsg/

Forfurther descriptions of individual resources, see main text. HGNC: the HUGO (human genome organization) Gene Nomenclature Committee; NCBI: the US National Center for Biotechnology Information; UCSC: University of California at Santa Cruz.

Resource use	Popular resources	Website address
Identifying related sequences and homologs	BLAST programs BLAT HomoloGene HCOP (orthology predictions)	http://blast.ncbi.nlm.nih.gov/Blast.cgi https://genome.ucsc.edu/cgi-bin/hgBlat?command=start http://www.ncbi.nlm.nih.gov/homologene http://www.genenames.org/tools/hcop
Protein sequence analysis	UniProt InterPro	http://www.uniprot.org https://www.ebi.ac.uk/interpro/
Genome browsers	Ensembl UCSC Genome Browser	http://www.ensembl.org/ https://genome.ucsc.edu/
Genome annotation	GENCODE NCBI Annotation (release 109)	https://www.gencodegenes.org/ https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Homo_sapiens/109/

Forfurther descriptions of individual resources, see main text. HGNC: the HUGO (human genome organization) Gene Nomenclature Committee; NCBI: the US National Center for Biotechnology Information; UCSC: University of California at Santa Cruz.

Specialized electronic gateways with electronic links to numerous such electronic resources are especially useful. If the starting point of interest is a gene, gene product or genetic disorder, the HGNC portal organized by the HUGO Gene Nomenclature Committee has a simple, user-friendly architecture. A comprehensive coverage of electronic resources is also available through the Human Genome Resources section of the US National Center for Biotechnology Information (NCBI)—see Table 2.3).

Gene nomenclature and the HGNC gateway

Gene symbols for human genes are allocated by the HUGO Gene Nomenclature Committee (HGNC). They typically have between three and seven characters, and are displayed in italicized uppercase, such as *HBB* (hemoglobin beta subunit), *CFTR* (cystic fibrosis transmembrane regulator), or *RN7SL* (7SL RNA). Mitochondrial genes are prefixed by *MT-* (for example, *MT-RNR1* is the mitochondrial 12S rRNA gene). Pseudogenes, naturally occurring but defective copies of a normally functional gene, usually have a symbol that is the same as a related functional gene, but followed by a P, or by a P followed by a number (for example, *CFTRP3* is one of three pseudogenes related to the *CFTR* gene). Note that the format for gene symbols for other species is often different. For example, the mouse and rat orthologs of the human *CFTR* gene are each given the symbol *Cftr*.

The HGNC portal at www.genenames.org has links to many databases and browsers. It can be interrogated by using as a query a gene symbol, if known, or descriptive text for an associated gene product or disease. For example, entering hemoglobin as the search term returns a list of genes

encoding the different subunits of all forms of human hemoglobin, and entering cystic fibrosis yields related results including the cystic fibrosis gene *CFTR*. Selecting a gene symbol such as *CFTR* opens the way to an extraordinary amount of related information through linked databases and genome browsers (Figure 2.10). It can also be used to identify groups of related genes under the section entitled Gene groups.

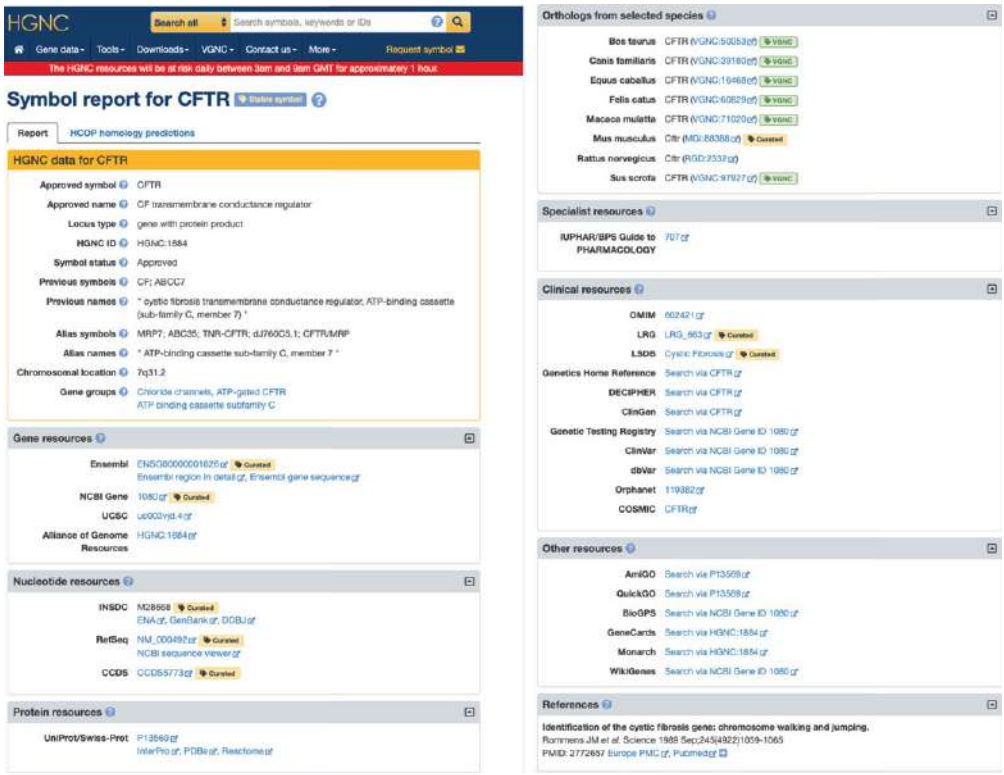


Figure 2.10 Getting a wealth of information on a selected human gene starting from the HGNC portal at www.genenames.org. The figure shows the output displayed after using *CFTR* (the human cystic fibrosis gene) as a search query. Under the Report tab, the box at top left outlined in yellow shows basic items of information maintained by HGNC, including Gene groups at bottom. The other sections, with titles highlighted in gray, provide links to numerous external databases, with information on associated gene, RNA and protein sequences, databases with clinical information and mutant sequences, and information on closely related orthologs in other species. Pressing on the second tab at top left, marked, HCOP homology predictions, identifies related sequences in a very wide range of organisms.

Databases storing nucleotide and protein sequences

In the example of Figure 2.10, the link for "Nucleotide resources" begins with *gen- eral* nucleotide sequence databases that are part of the International Nucleotide Sequence Database Collaboration (INSDC), comprising the European Nucleotide Archive (ENA), GenBank and the DNA Database of Japan (DDBJ). The ID number for the *CFTR* mRNA sequence in these databases is M28668, and the 6129-nucleotide sequence is presented. Under "Protein resources" is the useful UniProt/Swiss-Prot database where extensive information on the CFTR protein (ID number: P13569) can be found.

The problem with general nucleotide and protein sequence databases is that they contain *redundant* sequences (sometimes with many entries for the same sequence from independent DNA clones, some having partial sequences, some full length). To make it much easier to find a complete sequence of

interest, the RefSeq databases were established at the NCBI to provide a comprehensive, nonredundant, and well-annotated set of reference sequences for different species. The standard RefSeq database has non-redundant reference sequences for mRNAs (ID numbers prefixed by NM_); noncoding RNAs (ID numbers prefixed by NR_); and proteins inferred from a mRNA (ID numbers prefixed by NP_). The separate RefSeqGene database stores gene reference sequences.

Finding related nucleotide and protein sequences

Sequences evolutionarily related to a query nucleic acid or protein sequence are most often identified using one of the BLAST programs hosted at the NCBI. BLASTN uses a nucleotide query to compare with other nucleotide sequences, and BLASTP compares a protein sequence against other protein sequences (see Figure 2.9 for an example output). TBLASTN is powerful because it can compare a protein sequence against all possible translations of all sequences in nucleotide databases. Significant homology may be apparent across the length of the query sequence or be limited to a region, such as a conserved protein domain or some other shared sequence.

The BLAT program allows rapid sequence searching across whole genomes. Query nucleotide sequences (up to a total of 25 000 nucleotides) and query protein sequences (up to 10 000 amino acids) can be entered to search for homologous sequences across the human genome, or the genome of any of multiple model organisms. The output lists significant hits, with given chromosome coordinates and sequence alignments.

The sequence comparisons include searching to find equivalent sequences (*orthologs*) in other species, using programs such as Homologene and HCOP. The HGNC reports for genes show multiple orthologs from other species, such as shown in the *CFTR* gene report in Figure 2.10.

Links to clinical databases

As shown in Figure 2.10 HGNC reports also provide links to a variety of clinical databases under the section entitled “Clinical resources”. They include the On-line Mendelian Inheritance in Man (OMIM), Genetic Home Reference and various databases recording disease-associated mutations, including COSMIC (focusing on cancer) and ClinVar (which documents relationships between human DNA variants and phenotypes).

2.5 THE ORGANIZATION AND EVOLUTION OF THE HUMAN GENOME

Our genome has some curious characteristics, such as: division into a massive complex genome in the nucleus and a tiny simple genome in our mitochondria; a huge number of repetitive DNA sequences; a profusion of pseudogenes; a vast range in gene sizes; and a small proportion of functionally important nucleotides. In order to make sense of how our genome is organized, we begin this section by examining important evolutionary forces that shaped the genome.

A brief overview of the evolutionary mechanisms that shaped our genome

The widely accepted endosymbiont hypothesis proposes that our two physically separated genomes, the nuclear and mitochondrial genomes, originated when a type of aerobic prokaryotic cell was endocytosed (engulfed) by an anaerobic eukaryotic precursor cell, at a distant time in the past when

oxygen started to accumulate in significant quantities in the Earth's atmosphere. Over a long period, much of the original prokaryote genome was excised, causing a large decrease in its size, and this much reduced prokaryotic genome gave rise to the mitochondrial genome. DNA fragments excised from the aerobic prokaryote cell were transferred to the genome of the engulfing cell. The latter genome increased in size as a result, but then progressively went on to undergo further very significant changes in both size and form during evolution, developing into our nuclear genome.

The theory explains why mitochondria have their own ribosomes and their own protein-synthesizing machinery and why our mitochondrial DNA closely resembles in form a reduced (stripped-down) bacterial genome. But how did the genome of the engulfing cell become so large and complex? That largely happened by a series of different mechanisms that copied existing DNA sequences and added them to the genome. After some considerable time, the copies can acquire mutations that make them different from the parent sequences; ultimately new genes, new exons, and so on can be formed in this way.

Whole genome duplication is a quick way of increasing genome size, and comparative genomics has provided very strong evidence that this mechanism has occurred from time to time in different evolutionary lineages. There is compelling evidence, for example, that whole genome duplication occurred in the early evolution of our chordate ancestors just before the appearance of vertebrates.

Additional duplications of moderately large to small regions of DNA occur comparatively frequently on an evolutionary timescale, and they also give rise ultimately to novel genes, novel exons, and so on. They occur by copying mechanisms that work at the level of genomic DNA, or at the RNA level by using *reverse transcriptases* (RNA-dependent DNA polymerases) to make DNA copies of RNA transcripts that then insert into the genome.

Comparative genomics can reveal when new genes were formed in evolution by screening the genomes of multiple species to identify those that possess versions of the same gene (the different versions, such as the human *CFTR*gene and the mouse *Cftr* gene are said to be *orthologs*). In the examples given in Section 2.3, the gene encoding the p53 tumor suppressor first appeared at the time of bony vertebrates (it is not present in invertebrates or in non-bony vertebrates), but others have evolved very recently (the *TCP10L* gene, for example, is found only in Old World monkeys, apes, and humans.

The duplication mechanisms that led to a progressive increase in genome size and to the formation of novel genes and other novel functional sequences are, to a limited extent, offset by occasional evolutionary loss of functional DNA sequences, including genes. After whole genome duplication, for example, many of the new gene copies pick up mutations that cause them to be silenced, and they are eventually lost. And the Y chromosome is believed to have shed many genes over hundreds of million years. Gene loss can happen on a smaller scale, too.

Gene birth and gene loss are comparatively infrequent events, and even though humans and mice diverged from a common evolutionary ancestor about 80 million years ago, our gene repertoire is extremely similar to that of the mouse. However, *cis*-acting regulatory sequences such as enhancers often evolve rapidly, and although we have much the same set of genes as a mouse, they are often expressed in different ways. Differential gene regulation is a primary explanation for the differences between species that are evolutionarily closely related.

How much of our genome is functionally significant?

We will end this chapter by looking in some detail at different facets of our genome. But first, let us step back and take a broad look at its design. Here is one perspective attributed to the evolutionary biologist David Penny: "I would be quite proud to have served on the committee that designed the *[Escherichia] coli* genome. There is, however, no way that I would admit to serving on a committee that designed the human genome. Not even a university committee could botch something that badly."

The *E. coli* genome is a sleek genome, packed with gene sequences (90 % of the genome is made up of coding DNA sequences). By contrast, like the genome of many complex organisms, our genome seems rather flabby: coding DNA accounts for just 1.2 % of the genome, and the majority of the DNA is made up of highly repetitive noncoding DNA sequences of questionable functional value (see **Figure 2.11**). For decades much of our genome had largely been regarded as "junk DNA," an idea supported by the lack of correlation between genome size and organism complexity (the genome of the diploid onion, for example, is more than five times the size of our genome).

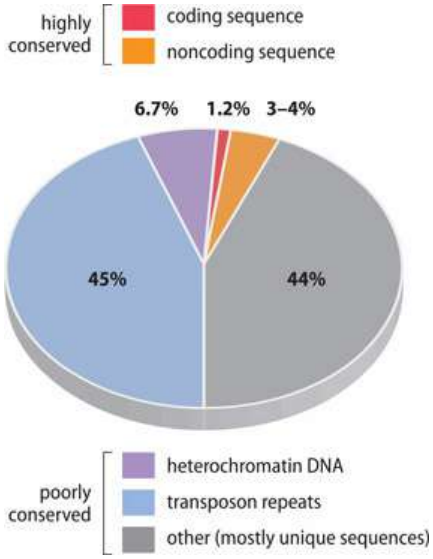


Figure 2.11 Human genome organization: extent of evolutionary conservation and repetitive sequences. Only just over 1.2 % of our genome is coding DNA that specifies protein sequences, and another roughly 3–4 % or so of our genome is made up of noncoding DNA sequences that have been highly or moderately conserved during evolution (as determined by looking at nucleotide substitutions in mammalian sequence alignments). Some of this conserved sequence is present in multiple copies and includes different types of repeated genes (gene families). The 6.7 % of our genome that is located in constitutive heterochromatin is very largely made up of poorly conserved repetitive DNA sequences that include sequences responsible for centromere function. Transposon repeats include highly repetitive interspersed repeats such as Alu and LINE-1 repeats; it is thought that during evolution some of these repeats contributed to the formation of new exons and regulatory elements, including some long ncRNAs.

The ENCODE Project seemed to offer a different perspective. The data suggested that much of the genome was transcribed, often from both DNA strands in specific regions, and 80.4 % of the human genome was claimed to participate in at least one RNA-associated or chromatin-structure-associated event in at least one cell type. However, the possible conclusion that much of our genome might be functionally significant has been strongly resisted by evolutionary biologists. Part of the difficulty in interpreting the ENCODE data is that much of the 80.4 % figure comes from the observed

representation of RNA transcripts, but many RNAs are produced at very low levels and their functional status is uncertain.

Sequence conservation due to selection and estimating functional constraint

Although mutation can potentially change any nucleotide in DNA, there must be constraints on changing a functionally important sequence during evolution. If we take a protein-coding sequence, for example, even a single amino acid change might quite often result in loss of the protein's function or produce an aberrant protein that might contribute to disease.

Mutations that result in adverse changes to the phenotype are effectively removed from populations over generations. That happens by a type of Darwinian natural selection called **purifying selection.** Compared with normal alleles, the mutant allele is not efficiently transmitted to subsequent generations (some people that carry it will not reproduce as well as people with normal alleles). Over long periods of evolutionary time, therefore, protein-coding sequences are constrained by the need to maintain function (*func- tional constraint*), and they change slowly in comparison with most other DNA sequences.

The amount of the human genome that is highly conserved (under functional constraint as a result of purifying selection) was initially estimated to be about 5 % (see Figure 2.11). However, that figure came from comparisons with many different mammals. Additional functional constraint became apparent in the 1000 Genomes Project, in which multiple human genomes were compared.

Functionally important DNA sequences that are rapidly evolving might not be seen to be conserved (and therefore constrained) when comparisons are made between a broad range of mammalian species. Although some rare non-coding DNA sequences are very strongly conserved—sometimes more than coding DNA—it is clear that many regulatory DNA sequences and RNA genes are rapidly evolving. They do, however, make important contributions to functional constraint in narrower evolutionary lineages, including primate and then human lineages in our case.

Taking that into account, the proportion of our genome that is subject to purifying selection is now thought to be of the order of at most 10 %; on that basis, most of our genome does not seem to have a valuable function. However, there is evolutionary value in having a large genome with surplus non-functional DNA because the non-functional DNA component can, through successive mutations, provide new functional sequences in the future, as described below.

The mitochondrial genome: economical usage but limited autonomy

David Penny's comment about the human genome in general (see above) certainly does not apply to the mitochondrial genome. Our mitochondrial DNA closely resembles in form a reduced (stripped-down) bacterial genome. Mitochondria, like chloroplasts, have their own ribosomes and their own protein-synthesizing machinery and almost certainly originated when a prokaryotic cell was engulfed by an anaerobic eukaryote precursor cell, allowing aerobic eukaryotes to develop.

The human mitochondrial genome has a total of 37 genes. Of these, 24 are RNA genes that make all the RNA required for protein synthesis in the mitochondrial ribosomes: the two rRNAs and 22 tRNAs (**Figure 2.12**). The remaining genes make 13 out of the 89 polypeptide subunits of the oxidative phosphorylation system (OXPHOS). (The other 76 OXPHOS subunits, like all other

mitochondrial proteins, are encoded by nuclear genes and synthesized on cytoplasmic ribosomes before being imported into the mitochondria.)

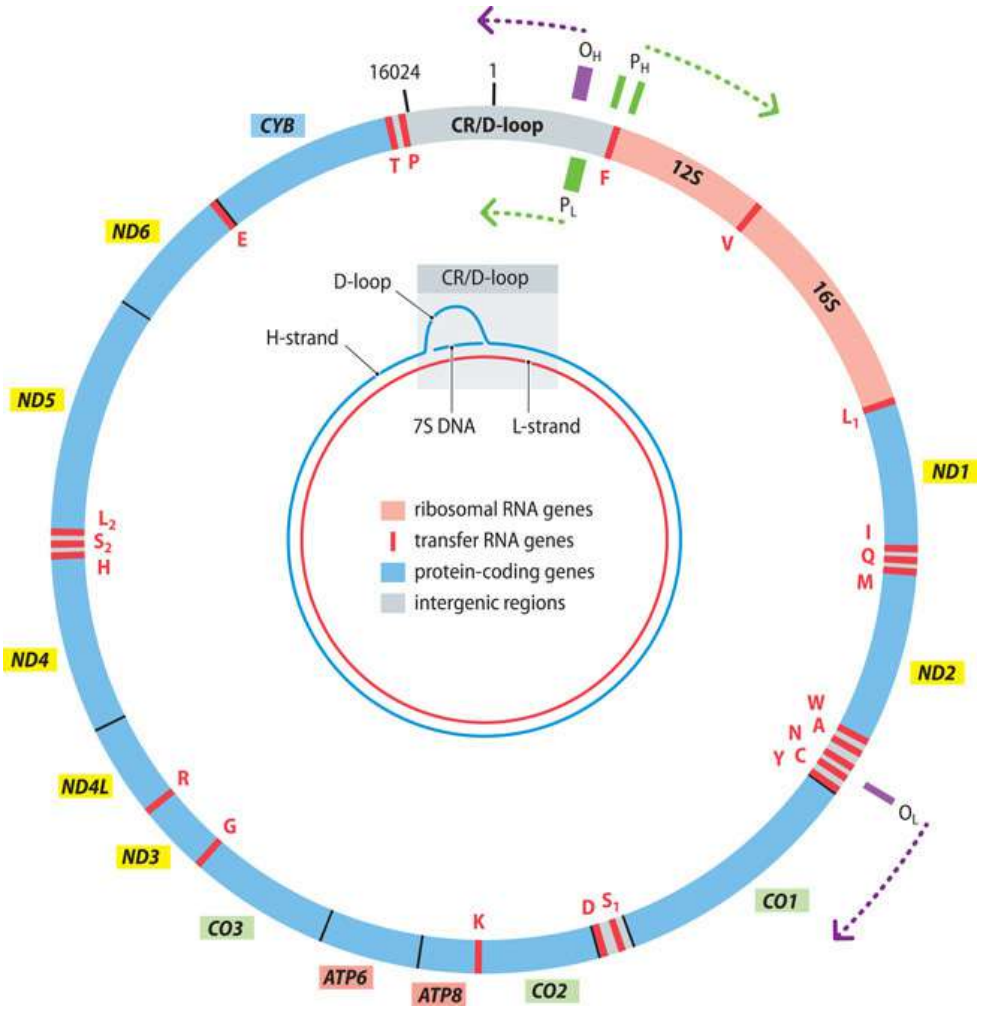


Figure 2.12 Organization of the human mitochondrial genome. The circular 16.6 kb genome has a heavy (H)-strand rich in guanines and a light (L)-strand rich in cytosines. 24 RNA genes, two making the 12S and 16S rRNAs and 22 making tRNAs (shown as thin red bars with a letter corresponding to the amino acid specified). The 13 protein-coding genes make a few components of the oxidative phosphorylation system: seven NADH dehydrogenase subunits (encoded by *ND1*-*ND6* and *ND4L*), two ATP synthase subunits (encoded by *ATP6* and *ATP8*), three cytochrome *c* oxidase subunits (encoded by *CO1*–*CO3*), and one cytochrome *b* (encoded by *CYB*). The reference nucleotide sequence begins from nucleotide 1 in the middle of the CR/D loop region and progresses in a clockwise direction (CR for control region because it has regulatory sequences; D-loop signifies a displacement loop causing a locally triple-stranded DNA to form after repeat synthesis of a short sequence called 7S DNA). Two promoters, P_H and P_L (green boxes), which transcribe respectively the heavy and light strands in opposite directions (clockwise and counterclockwise), generate large multigenic transcripts from each strand that are subsequently cleaved. For further information, see the MITOMAP database at http://www.mitomap.org. For the revised Cambridge human mtDNA reference sequence, see http://mitomap.org/MITOMAP/HumanMitoSeq

Because of the need to make just 13 different proteins, the genetic code used by mitochondrial DNA (MtDNA) has been allowed to drift a little from the “universal” genetic code that is used for

nuclear DNA. It uses four stop codons, for example. We consider the genetic code in some detail in Section 7.1; the differences between the nuclear and mitochondrial genetic codes are given in Figure 7.2.

None of the mitochondrial genes is interrupted by introns, and the genome is a model of economical DNA usage: close to 95 % of the genome (all except 1 kb out of the 16.6 kb of DNA) makes functional gene products. Note that transcription of the two DNA strands occurs using one promoter each to generate large multigenic transcripts that are subsequently cleaved to generate individual mRNAs and ncRNAs.

Gene distribution in the human genome

More than 90 % of the mitochondrial DNA sequence directly specifies a protein or functional ncRNA, and there is one intronless gene every 450 bp on average. The nuclear genome is very different: the gene density is much lower, genes are frequently interrupted by introns as listed in Table 2.1 above, and a sizable fraction of the genome is made up of repetitive DNA, notably highly repetitive non-coding DNA.

Close to 7 % of the nuclear genome is located in constitutive **heterochromatin** that remains highly condensed throughout the cell cycle (the chromosomal locations of human heterochromatin are given in Figure 2.8 above). Although almost devoid of genes, the constitutive heterochromatin of somatic cells does contain a tiny number of genes that are inactive in somatic tissues but expressed in germ cells (where chromatin has a more open structure). The remaining 93 % of our genome is accommodated in less-condensed, gene-rich **euchromatin**.

Protein-coding genes have been comparatively easy to identify. GENCODE data estimate close to 20 000 human protein-coding genes, but the number can never be exact because of variation between individuals (and sometimes between maternally and paternally inherited genomes) in copy number for some repeated genes).

Identifying RNA genes is much more problematic. Three characteristics make it difficult to do so: the lack of a sizable open reading frame, lack of evolutionary sequence conservation, and in some cases extremely small sizes (making them easily overlooked). Establishing the functional significance of many transcribed noncoding DNA sequences can therefore be difficult. The most recent GENCODE data identify more RNA genes than protein-coding genes (**Table 2.4**), but the functional status of some putative RNA genes remains unproven.

TABLE 2.4 A SNAPSHOT OF THE NUMBERS OF HUMAN GENES AND PSEUDOGENES LISTED BY GENCODE VERSION 40 (RELEASED IN APRIL 2022)

Class	Number
PROTEIN-CODING GENES	19988
RNA GENES	26372
making long ncRNA	18805
making short ncRNA	7567
PSEUDOGENES	14774
processed	10661

Obtained at http//gencodegenes.org/human/stats.html

Class	Number
unprocessed	3566
other	547

Obtained at http//gencodegenes.org/human/stats.html

Genome sequencing showed that gene and exon density in the euchromatic regions can vary enormously. Some chromosomes are gene-rich, such as chromosomes 19 and 22; others are gene-poor, notably the Y chromosome (which makes only 31 different proteins that mostly function in male determination). Within a chromosome, the pattern of alternating dark and light bands reflects different base compositions, and also differences in gene and exon density (as described in the legend to Figure 2.8).

The extent of repetitive DNA in the human genome

Our large nuclear genome is the outcome of periodic changes that have occurred over very long timescales during evolution, including rare whole genome duplication and intermittent chromosome rearrangements, localized DNA duplications, DNA duplication followed by dispersal to other genome locations, and loss of DNA sequences. The net result has been a gradual increase in DNA content and gene number through evolution.

Previous whole genome duplications were followed by a gradual loss of most of the duplicated sequences; accordingly, unique sequences make up quite a sizable fraction of our genome. Nevertheless, highly repetitive DNA sequences originating from transposons (mobile DNA elements; see below) plus the repetitive DNA families found in heterochromatin account for more than 50 % of our genome (described in Figure 2.11 above).

In addition to transposon-derived repeats, our euchromatin contains clear evidence of localized DNA duplications. In some cases, the repeats have diverged considerably in sequence—the duplication occurred many tens or hundreds of millions of years ago in evolution, and subsequent mutations have led to divergence in sequence between the repeats. But other localized duplications are quite striking because they have occurred very recently in evolution. For example, about 5 % of our euchromatin DNA consists of neighboring duplicated segments that are more than 1 kb long and show more than 90 % sequence identity. Many of these **segmental duplications** are primate-specific, and they are particularly common close to telomeres and centromeres (about 40 % occur in subtelomeric regions; about 33 % occur at pericentromeric regions).

There is a significant amount of repetitive coding DNA within genes and also many repeated genes. Within a gene, repetitive coding DNA may be found in an individual exon (usually as a tandem duplication of one or more nucleotides), or one or more exons has been repeated (**Figure 2.13A** and **2.13B**).

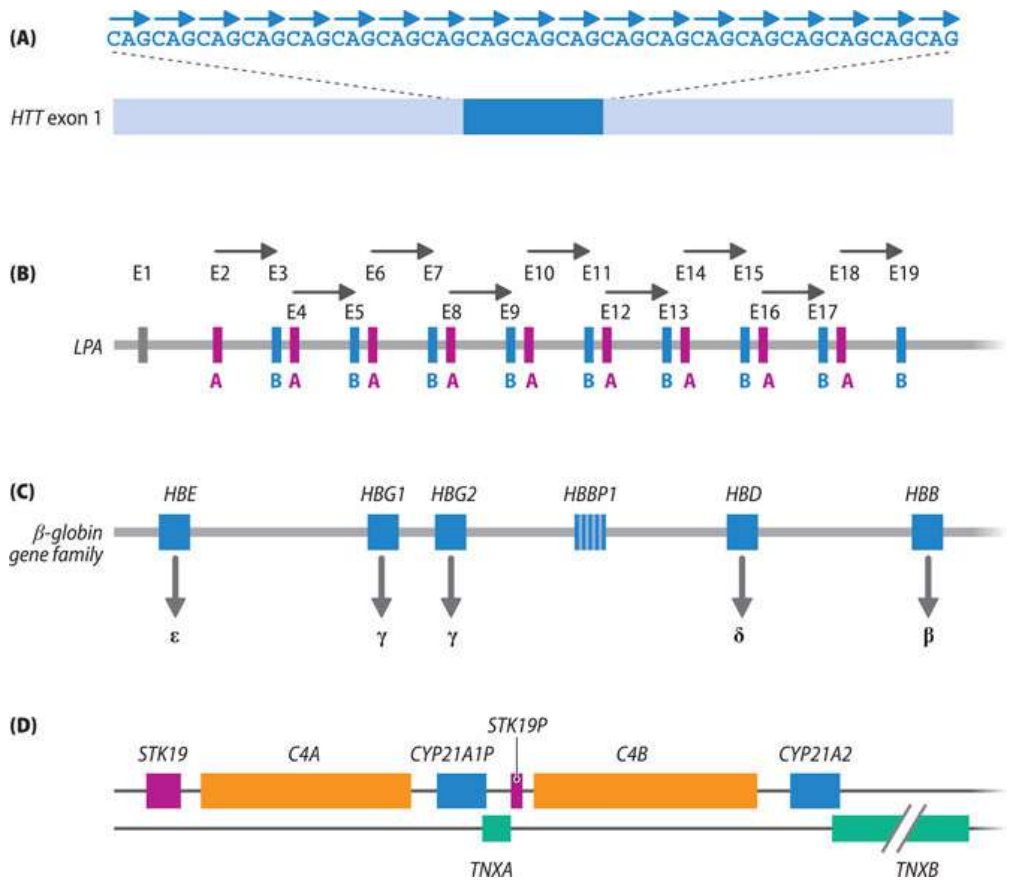


Figure 2.13 Examples of tandemly repetitive coding DNA and clustered gene families. (A) Normal alleles of the *HTT* huntingtin gene have an array of tandemly repeated CAG codons in exon 1 that varies in number up to 35 repeats (having more than 36 repeats results in Huntington disease). (B) The *LPA* gene encodes lipoprotein Lp(a), a protein with multiple kringle domains that are each 114 amino acids long and extremely similar in sequence. Each kringle repeat is encoded by a tandemly repeated pair of exons (here labeled A and B) that encode two adjoining parts of the kringle domain and that can be present in different copy numbers. The example shown here has nine pairs of A and B exons, starting with exons 2 and 3 (E2 and E3) and continuing through to exons 18 and 19 (E18 and E19). (C) The b-globin gene family has six highly related genes. Four genes make alternative globins used in hemoglobin, but the status of *HBD* is uncertain (q-globin is never incorporated into a hemoglobin protein) and *HBBP1* is a pseudogene. (D) The HLA (human leukocyte antigen) region of normal individuals has a tandemly duplicated unit containing four gene sequences, encoding serine threonine kinase 19, complement C4, cytochrome P450 21-hydroxylase, and tenascin-X (transcribed from the opposite strand). Subsequently, three of the genes became pseudogenes, having acquired inactivating mutations *(CYP21A1P)* or having also lost significant amounts of sequence *(STK19P* and *TNXA)*.

On a larger scale, the repeated unit can consist of a whole gene or occasionally two or more unrelated genes (Figure 2.13C and 2.13D). The resulting multi-gene families contain two or more genes that produce related or even identical gene products. The more recently duplicated genes are readily apparent because they make very closely related or identical products. Genes originating from more evolutionarily ancient duplications make more distantly related products.

The organization of gene families

Different classes of multigene families exist in the human genome, and the number of genes in a gene family can range from two to many hundreds (**Table 2.5**). Some are clustered genes confined to one subchromosomal region. They typically arise by tandem gene duplication events in which chromatids first pair up unequally so that they become aligned out of register over short regions. The mispaired chromatids then exchange segments at a common breakpoint (**Figure 2.14**).

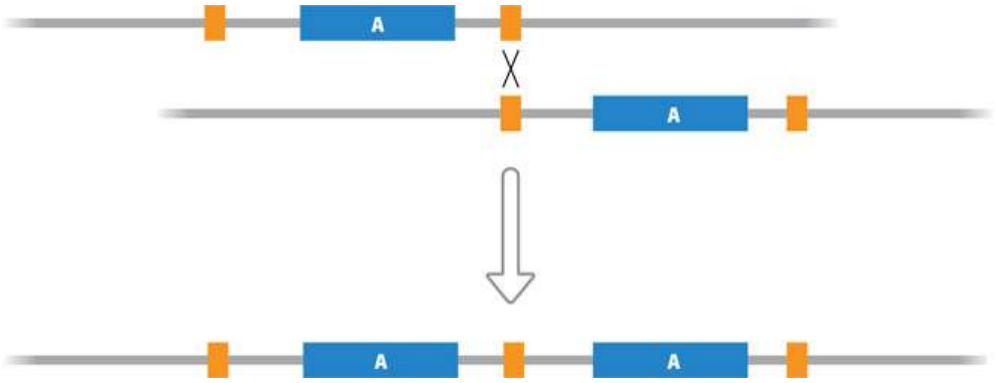


Figure 2.14 Tandem gene duplication. Gene duplication can occur after sister chromatids or non-sister chromatids of homologous chromosomes mispair so that they are slightly out of alignment. In this example, the result is that gene A on one chromatid is out of register with gene A on the opposing chromatid. Subsequent breakage of both chromatids at the position marked by the cross (X) and swapping of fragments between chromatids can result in a chromatid that has two copies of gene A (the left fragment of the top chromatid joins to the right fragment of the bottom chromatid). The exchange may be facilitated by base pairing between highly related noncoding repetitive DNA sequences (orange boxes).

TABLE 2.5 EXAMPLES OF MULTIGENE FAMILIES IN THE HUMAN GENOME

Gene family	Copy number	Genome organization
β-Globin	6 (includes one pseudogene)	clustered within 50 kb at chromosome 11 p15 (Figure 2.13C)
Class I human leukocyte antigen (HLA)	17 (includes many pseudogenes and gene fragments)	clustered over 1.8Mb at 6p21.3
Neurofibromatosis type I	1 functional gene; 8 unprocessed pseudogenes	functional gene, NF1, at 17q11.2; pseudogenes dispersed over pericentromeric regions on several other chromosomes
Ferritin heavy chain	1 functional gene; 27 processed pseudogenes[a]	functional gene, FTH1, at 11q13; pseudogenes dispersed over multiple chromosome locations
U6 snRNA	49 genes; 800 processed pseudogenes[a]	scattered on many chromosomes

a A processed pseudogene (retro pseudogene) is a copy of a gene transcript and so has counterparts of exon sequences only. By contrast, an unprocessed pseudogene (which is a copy of the genomic sequence) also has sequences corresponding to introns and upstream promoters); see Box 2.4.

The α- and β-globin gene clusters on chromosomes 16 and 11, respectively, arose by a series of tandem gene duplications. Some of the duplicated genes (such as the *HBA1* and *HBA2* genes, which make identical a-globins, or the *HBG1* and *HBG2* genes, which make g-globins that differ at a single

amino acid) are the outcome of very recent gene duplication. Other globin genes are clearly related to each other but have more divergent sequences. The different globin classes have slightly different properties, an advantage conferred by gene duplication (see below).

Other gene families are distributed over two or more different chromosomal regions. In some cases they originally arose from duplicated genes in gene clusters that were then separated by chromosome rearrangements. In other cases, cellular reverse transcriptases were used to make natural complementary DNA (cDNA) copies of the mRNA produced by a gene, and the cDNA copies were able to insert successfully elsewhere in the genome of germline cells. Because the cDNA copies of mRNA lacked promoters, as well as intron sequences, they very frequently degenerated into nonfunctional **pseudogenes**. For some genes, however, cDNA copies have integrated during evolution at other chromosomal locations to produce functional genes (**Box 2.4**).

BOX 2.4 PSEUDOGENES

One common consequence of gene duplication is that a gene copy diverges in sequence but instead of producing a variant gene product it gradually accumulates deleterious mutations to become a **pseudogene**. A pseudogene copy of a protein-coding gene can usually be detected by identifying deleterious mutations in the sequence that corresponds to the coding DNA sequence; RNA pseudogenes are less easy to identify as pseudogenes. GENCODE lists more than 14 000 pseudogenes in the human genome (Table 2.4). According to their origin, pseudogenes can be divided into two major classes: unprocessed pseudogenes and processed pseudogenes. Despite their name, some pseudogenes are known to have functionally important roles, as described below.

UNPROCESSED PSEUDOGENES

An unprocessed pseudogene arises from a gene copy made at the level of genomic DNA, for example after tandem gene duplication (Figure 2.14). Initially, the copied gene would have copies of all exons and introns of the parental gene plus neighboring regulatory sequences including any upstream promoter. Acquisition of deleterious mutations could lead to gene inactivation (‘silencing’) and subsequent decay, and sometimes instability (substantial amounts of the DNA sequence can be lost, leaving just a fragment of the parental gene). Unprocessed pseudogenes are typically found in the immediate chromosomal vicinity of the parental functional gene (see the example of *HBBP1* in Figure 2.13C). Sometimes, however, they are transposed to other locations because of instability of pericentromeric or subtelomeric regions. For example, the *NF1* neurofibromatosis type I gene is located at 17q11.2, and eight highly related unprocessed *NF1* pseudogenes are found (with one exception) in pericentromeric regions of other chromosomes as a result of comparatively frequent interchromosomal exchanges at pericentromeric regions.

PROCESSED PSEUDOGENES (RETROPSEUDOGENES)

A processed pseudogene arises by reverse transcription of an RNA from a parental gene followed by random integration of the resulting cDNA copy elsewhere in the genome (**Figure 1**). The cDNA copy lacks any sequences corresponding to introns and regulatory sequences occurring outside exons, such as upstream promoters. Integration of a cDNA copy of a protein-coding gene will usually mean that the cDNA is not expressed, and it will acquire deleterious mutations to become a

retropseudogene. If, however, the cDNA integrates at a position adjacent to an existing promoter it may be expressed and acquire some useful function to become a **retrogene** (see Figure 1).

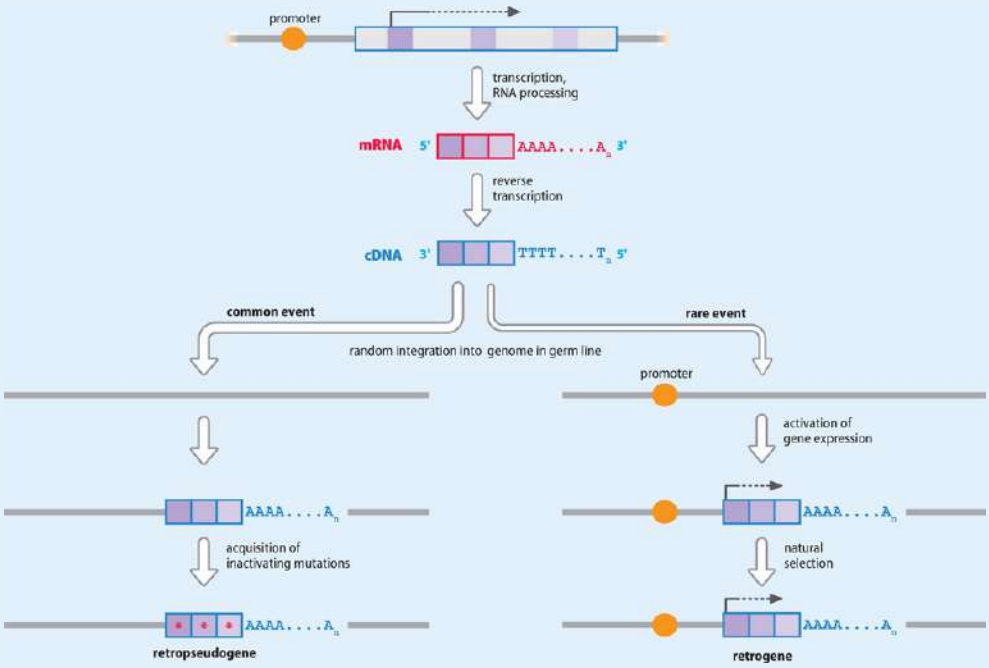


Figure 1 Retrogenes and retropseudogenes originate following reverse transcription from RNA transcripts. In this example, a protein-coding gene with three exons is transcribed from an upstream promoter, and introns are excised from the transcript to yield an mRNA. The mRNA can then be converted naturally into an antisense single-stranded cDNA by using a cellular reverse transcriptase. If this occurs in the germline and the resulting transcript then integrates randomly into the genome, it will most probably be incapable of expression (it will lack a promoter) and will degenerate into a retropseudogene, after acquiring harmful inactivating mutations (red asterisks, bottom left). If, however, the transcript inserts next to an endogenous promoter it may be expressed. Very occasionally, and according to need, the expressed gene may be useful and be preserved by natural selection as a functional retrogene (bottom right).

Several protein-coding genes have associated retropseudogenes (Table 2.5), but the highest numbers of retropseudogenes derive from RNA genes. RNA genes transcribed by RNA polymerase III have *internal pro- moters*. That is, the sequences needed to attract the transcriptional activation complexes (which then go on to recruit an RNA polymerase) are located within the transcription unit itself, instead of being located upstream like the promoters of protein-coding genes. When transcripts of these genes are copied into cDNA, the DNA copies of their transcripts also have promoter sequences, giving the potential to reach very high copy number. Human Alu repeats, for example, originated as cDNA copies of 7SL RNA, and the high copy number mouse B1 and B2 repeats are diverged cDNA copies of tRNAs.

FUNCTIONAL PSEUDOGENES

Comparative genomic studies indicate that some pseudogene sequences have evolved under purifying selection (they are more evolutionarily conserved than would have been expected for a functionless DNA sequence). Many pseudogenes are known to be transcribed, and there is good evidence that the transcripts of some pseudogenes have important regulatory roles. For example, the

PTEN gene (which is mutated in multiple advanced cancers) is located on chromosome 10 and is regulated by an RNA transcript from a closely related processed pseudogene *PTENP1* located on chromosome 9. As described in Chapter 6, *PTENP1* regulates cellular levels of *PTEN*, and thereby acts as a tumor suppressor).

Many RNA genes are also members of large gene families. For example, the short arms of chromosomes 13, 14, 15, 21, and 22 each have 30–40 tandem repeats of a 45 kb DNA sequence that specifies 28S, 18S, and 5.8S ribosomal RNA. During mitosis, the megabase-sized clusters of ribosomal DNA on the different chromosomes can pair up and exchange segments, a type of interchromosomal recombination.

Noncoding RNAs can also be reverse transcribed to give cDNA copies that can integrate elsewhere in the genome, like the mRNA-derived cDNAs in Figure 1 of Box 2.4. During evolution, cDNA copies of certain classes of RNA genes that have internal promoters and are transcribed by RNA polymerase III were particularly successful in integrating into the genome because the copies carried with them promoter sequences. In some cases this resulted in a huge increase in copy number, and as explained below it gave rise to the most commonly occurring repetitive DNA sequence in the human genome, the Alu repeat.

The significance of gene duplication and repetitive coding DNA

Over long periods of evolutionary time there seems to have been a relentless drive to duplicate DNA in complex genomes. That has meant that whole genes have been duplicated to give gene families as described above. Tandem duplication of exons (which occurs by the same mechanisms that produce tandem gene duplication) is also evident in about 10 % of human protein-coding genes. There are several advantages of DNA sequence duplication, as listed below.

Gene dosage

Duplication of genes can be advantageous simply because it allows more gene product to be made. Increased gene dosage is an advantage for genes that make products needed in very large amounts in cells—we have hundreds of virtually identical copies of genes that make individual ribosomal RNAs and individual histone proteins, for example. Exon duplication might also be an advantage when an exon (or group of exons) encodes a structural motif that can be repeated, allowing proteins such as collagens to extend the size of structural domains during evolution.

Novel genetic variants

Once a gene or exon has duplicated, there are initially two copies with identical sequences. When that happens, the constraints on changing the sequence imposed by Darwinian natural selection may be applied to one of the two sequences only. The other sequence is free from normal constraints to maintain the original function; it can diverge in sequence over many millions of years to produce a different but related genetic variant. Divergent exons allowed the formation of different but related protein domains and the possibility of alternative splicing to produce transcripts with different exon

combinations. Additionally, as described below, certain types of mobile element allow the copying of exons from one gene to another (exon shuffling) to produce novel combinations of exons.

Divergent genes produced by tandem gene duplication allow the production of variant but related proteins. The vertebrate globin superfamily provides illustrative examples. Over a period of 800 million years or so, a single ancestral globin gene gave rise to all existing globin genes by a series of periodic gene duplications. Early duplications led to diverged gene copies that ultimately came to be expressed in different cell types, producing globins that were adapted to work in blood (hemoglobins), in muscle (myoglobin), in the nervous system (neuroglobin), or in multiple cell types (cytoglobin).

More recent duplications in the a-and b-globin gene clusters (see Figure 2.13C for the latter) led to different varieties of hemoglobin being produced at different stages of development. Thus in early development, zeta (z)-globin is used in place of a-globin, while epsilon (e)-globin is used instead of b-globin in the embryonic period, and g-globin is used instead of b-globin in the fetal period. The globins incorporated into hemoglobin in the embryonic and fetal periods have been considered to be better adapted to the more hypoxic environment at these stages.

There are, however, disadvantages to DNA sequence duplication. One consequence of repetitive coding DNA and tandemly repeated gene sequences is that the repeated DNA sequences can be prone to genetic instability, causing disease in different ways. We examine this in detail in Chapter 7.

Highly repetitive noncoding DNA in the human genome

Just over half of the human genome is made up of highly repetitive noncoding DNA sequences, of which a minority (about 14 %) is found in constitutive heterochromatin (which accounts for a total of about 7 % of our DNA). Euchromatin accounts for about 93 % of our DNA, of which just under half is made up of highly repetitive noncoding DNA (accounting for 45 % of the total genome).

The repetitive noncoding DNA in heterochromatin is a mixture of repetitive DNA sequences that are found in both heterochromatin and euchromatin (see examples below) plus DNA repeats that are characteristic of heterochromatin. The latter include different satellite DNA families of highly repetitive tandem repeats. Satellite DNAs are common at centromeres and include: alphoid DNA, with a 171 bp a repeat unit (found at all human centromeres); a 68 bp b repeat unit at the centromeres of the acrocentric chromosomes plus chromosomes 1, 9, and Y; plus different other satellite DNAs with comparatively small repeat units.

Like most heterochromatin DNA, the DNA of centromeric heterochromatin is very poorly conserved between species. Telomeric heterochromatin is the exception. It is based on TTAGGG repeats (that extend over lengths of 5–15 kb at the chromosome ends); the TTAGGG telomere repeat sequence is conserved throughout vertebrates and is highly similar to the telomere repeats of many invertebrates and plants.

Transposon-derived repeats in the human genome

Different classes of highly repetitive DNA sequences occur in an interspersed fashion (rather than as tandem repeats) and are commonly found within genes (usually in introns, but sometimes in exons). They arose from *transposons*, mobile elements that were able to migrate from one location in the genome to another.

The vast majority of transposon-derived repeats in the human genome can no longer transpose and are considered "transposon fossils". They are either truncated, having lost key sequences, or have picked up inactivating mutations during evolution. As a result, only a very small number of human transposon repeats are now capable of transposing autonomously (but other transposon-derived repeats can sometimes transpose by hitching a ride when located physically close to an autonomously transposing repeat).

Only about 6 % of the highly repetitive interspersed repeats in euchromatin originated from DNA transposon families that transpose by a cut-and-paste mechanism. The great majority, accounting for at least 40 % of the genome, originated from **retrotransposons** that transpose through an RNA intermediate. Here an RNA is copied by cellular reverse transcriptases to make a cDNA copy that integrates elsewhere in the genome (the same principle as shown in the figure in Box 2.4). There are three major classes and one minor class of retrotransposon-derived repeats in the human genome, as listed below and in **Table 2.6**.

TABLE 2.6 HUMAN TRANSPOSON REPEAT CLASSES AND FAMILIES

Transposon repeats by origin	Repeat class	Full length	% of genome	Examples (*see Fig. 2.15)
RETROTRANSPOSON REPEATS (via RNA intermediate)	LINEs	6–8 kb	21%	LINE-1*
	SINEs	100–300bp	13%	Alu repeat*
	Retrovirus-like elements	1.5–11kb	8%	HERV family
				LTR element
	SVA elements	~2kb	0.1 %	SVA element*
DNA TRANSPOSON REPEATS (via cut and paste)	Various	2–3 kb	3%	

Note: the great majority of repeats are truncated, having lost sequence components during evolution. Abbrevations are: LINEs—Long Interspersed Nuclear Elements; SINEs—Short Interspersed Nuclear Elements; HERV—human endogenous retrovirus; LTR—long terminal repeat; SVA—Sine-R-VNTR-Alu.

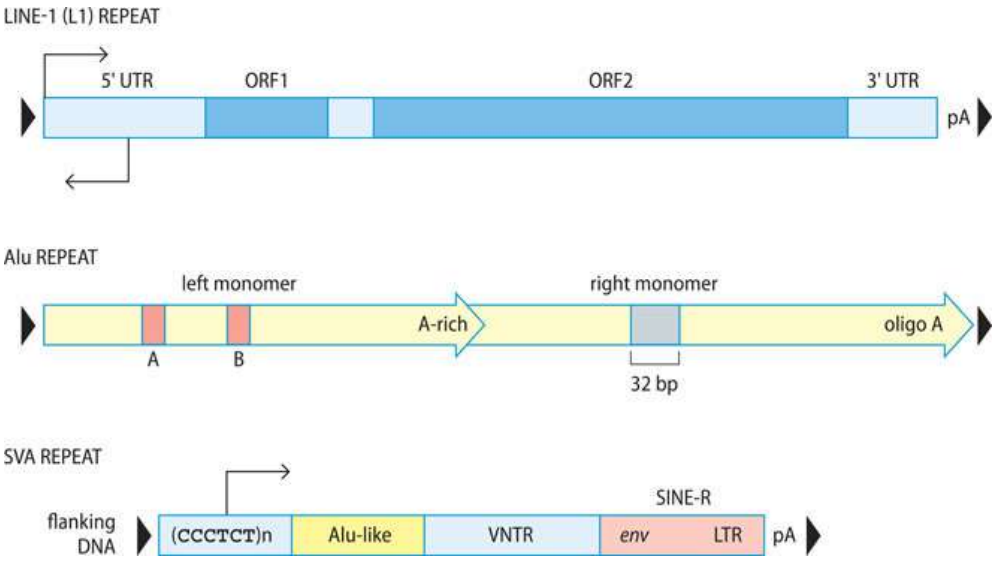


Figure 2.15 Structure of three types of commonly transposing human transposon-derived repeats. Some full-length LINE-1 repeats can transpose autonomously. The ORF2 open reading frame makes an endonuclease that can cut DNA (preferentially at AT-rich sequences) and a reverse transcriptase that uses the released 3¢-OH end to prime cDNA synthesis. New insertion sites are flanked by a small target-site duplication (flanking black arrowheads). Alu repeats often consist of two monomer repeats with similar sequences terminating in an A-rich or oligo A sequences. Nonautonomous SVA repeats have both an Alu-like sequence and a 3¢ HERV fragment (SINE-R), separated by a sequence with variable number of tandem repeats (VNTR). They may often be transcribed from promoters in flanking DNA sequence. Arrows indicate propensity for transcription.

- *LINEs (long interspersed nuclear elements)*. Of the three LINE families, the most numerous is the LINE-1 (also called L1) family. The only human LINE elements currently capable of transposition are a small subset (about 80–100 copies) of full-length LINE-1 repeats.
- *SINEs (short interspersed nuclear elements)*. SINEs are non-autonomous: they need a reverse transcriptase to be supplied (for example by a neighboring LINE-1 repeat). The primate-specific Alu repeat family, with close to 1.5 million copies, is the dominant family and is preferentially located in euchromatic regions of the genome (unlike LINEs that tend to be located in heterochromatin). Alu repeats have evolved from cDNA copies of 7SL RNA (a component of the signal recognition particle that regulates transport of proteins out of cells). The other two SINE families have evolved respectively from reverse transcription of a tRNA and a 5S rRNA.
- *Retrovirus-like LTR elements.* The human endogenous retroviruses (HERVs) contain sequences resembling the key retroviral genes, *gag, pol* (encoding reverse transcriptase) and *env*, flanked by *long terminal repeats (LTRs),* but there is little evidence of actively transposing human HERVs. Truncated versions with remnants of the *gag* gene are also common.
- *SVA (**S** INE-R-**V** NTR-**A** lu) elements.* This very small family has compound repeats with a mix of an Alu-like sequence and a HERV-like sequence (confusingly known as SINE-R)—see **Figure 2.15**.

The evolutionary value of transposon repeats

At this stage, one might wonder why so much—at least 50 %—of our genome is composed of transposon-derived repeats, the great majority of which are now incapable of transposition. A likely answer is the drive to increase genetic novelty. In complex genomes a relentless evolutionary drive to duplicate DNA sequences has resulted in duplicated genes, duplicated regulatory sequences, duplicated exons, and so on, allowing the freedom to alter duplicates, while conserving the original functions of duplicated elements. Transposition can carry regulatory elements, and even exons, to different parts of the genome where they can change the functions of genes (for an example, see **Figure 2.16** for how a LINE-1 repeat can cause *exon shuffling* between genes). In very rare cases, transposable elements even appear to have given rise to new genes in evolution.

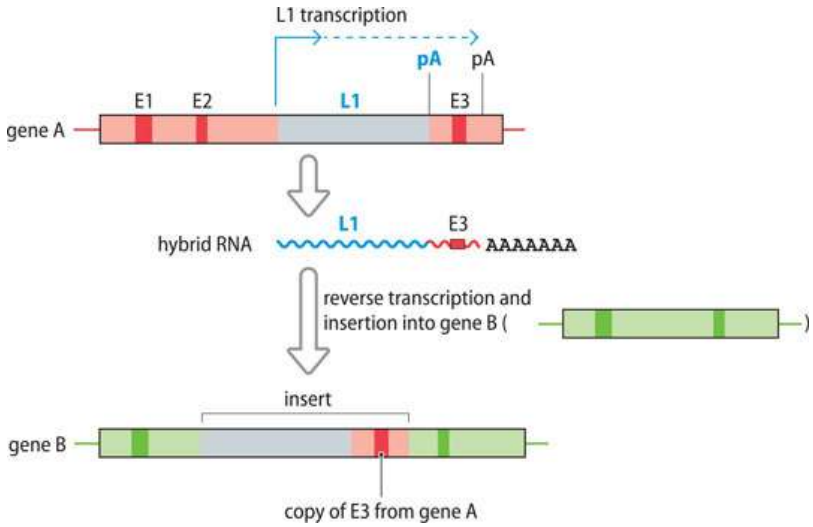


Figure 2.16 Retrotransposons can mediate exon shuffling. Exon shuffling can be carried out using retrotransposons such as actively transposing members of the LINE1 (L1) sequence family, as shown here. LINE-1 elements have weak poly(A) signals (pA) and so transcription often continues past such a signal until another downstream poly(A) signal is reached (for example after exon 3 (E3) in gene A). The resulting RNA copy contains a transcript not just of L1 sequences but also of a downstream exon (in this case E3). The L1 reverse transcriptase machinery can then act on the extended poly(A) sequence to produce a hybrid cDNA copy that contains both L1 and E3 sequences. Subsequent transposition into a new chromosomal location may lead to the insertion of exon 3 into a different gene (gene B)—Figure 1 Box 9.3.

SUMMARY

- Genes are transcribed to make RNA. Protein-coding genes make an mRNA that is decoded to make a poly-peptide. RNA genes make a functional noncoding RNA.
- An mRNA contains a central coding sequence, flanked by noncoding untranslated regions that contain regulatory sequences.
- In eukaryotes, the DNA sequence corresponding to the coding sequence of an mRNA is often divided into exons that are separated by intervening introns.
- The primary RNA transcript, an RNA copy of both exons and introns, is cleaved at exon-intron boundaries. Most of the transcribed intron sequences are discarded, and transcribed exon sequences are spliced together. Specialized end sequences—a 5¢ cap sequence and a 3¢ poly(A) sequence—protect the ends of the mRNA and assist in transfer to the cytoplasm to engage with ribosomes.
- The coding sequence of an mRNA is translated using groups of three nucleotides (codons) to specify individual amino acids that are bonded together to make a polypeptide.
- Introns are also found in noncoding DNA, both in some RNA genes and in many DNA sequences that make untranslated regions in mRNAs.
- Noncoding RNAs perform many different functions in cells, but most regulate gene expression, either ubiquitously (to assist general protein synthesis), or by controlling certain target genes in selected cell types.

- The human nuclear genome is composed of 24 different types of very long linear DNA molecules, one each for the 24 different types of human chromosomes (1–22, X, and Y). It has about 20 000 protein-coding genes and over 26 000 RNA genes (genes that make noncoding RNAs).
- Our mitochondrial genome consists of one type of small circular DNA molecule that is present in many copies per cell. It has 37 genes that make all the rRNAs and tRNAs needed for protein synthesis on mitochondrial ribosomes plus a few of the proteins involved in oxidative phosphorylation.
- The great majority of the genome consists of poorly conserved DNA sequences and only about 10 % of genome sequences are thought to be under selective constraint to maintain function.
- About 1.2 % of the nuclear genome is decoded to make proteins. These coding DNA sequences have mostly been highly conserved during evolution—for each human protein, recognizably similar proteins exist in many other organisms.
- RNA genes are more rapidly evolving than coding DNA, and while the shape of a noncoding RNA is particularly important the linear sequence is less important than that of polypeptides.
- Regulatory sequences are generally less conserved than coding DNA and polypeptide sequences. Humans and mice have a very similar set of genes but they are expressed differently because of differences in regulatory elements.
- Repetitive DNA sequences are very common in the human genome. They include both tandem repeats (often sequential head-to-tail repeats) and dispersed repeats.
- Tandem repeats may be found within genes and coding sequences, and whole genes can be duplicated several times to produce a clustered gene family. Other gene families are made up of gene copies that are dispersed across two or more chromosomes.
- Gene families often contain defective gene copies (pseudogenes and gene fragments) in addition to functional genes.
- Dispersed gene copies often arise in evolution from RNA transcripts that are copied by a reverse transcriptase to make a complementary DNA that integrates randomly into chromosomal DNA (retrotransposition).
- DNA sequence lying outside exons is largely composed of repetitive sequences, including highly repetitive interspersed repeats such as Alu repeats. They originated by retrotransposition (DNA copies were made of RNA transcripts that then integrated into the genome). Very few of the repeats are currently able to transpose.
- The DNA of centromeres and telomeres is largely composed of very many tandemly repeated copies of short sequences.
- Gene and exon duplication has been a driving force during genome evolution. Novel genes and exons are occasionally produced by tandem duplication events. Novel exons and novel regulatory sequences can also be formed by retrotransposition.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

More detailed treatment of much of the subject matter in this chapter, including a detailed account of human genome organization, gene evolution, and the Human Genome Project, can be found in the following:

Strachan T & Read AP (2019) *Human Molecular Genetics*, 5th ed. Garland Science.

Protein-coding genes and protein structure

Agris PF (2007) tRNA's wobble decoding of the genome: 40 years of modification. *J Mol Biol* 366:1–13; PMID 17187822.

Piovesan A (2019) Human protein-coding genes and gene feature statistics in 2019. *BMC Res Notes*: 12: 315; PMID 31164174

Preiss T & Hentze MW (2003) Starting the protein synthesis machine: eukaryotic translation initiation. *BioEssays* 25:1201–1211; PMID 14635255.

Whitford D (2005) *Protein Structure and Function*. John Wiley & Sons.

RNA genes and regulatory RNA

Amaral PP (2008) The eukaryotic genome as an RNA machine. *Science* 319:1787–1789; PMID 18369136.

Ponting CP (2009) Evolution and functions of long noncoding RNAs. *Cell* 136:629–641; PMID 19239885

Human genome: analysis and internet resources

Djebali S (2012) Landscape of transcription in human cells. *Nature* 489:101–108; PMID 22955620.

ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74; PMID 22955616.

GENCODE Project statistics (with useful statistics on the numbers of human genes and transcripts) are available at http://www.gencodegenes.org/human/stats.html

Genome Browsers and Internet Resources – see Table 2.3.

User's Guide to the Human Genome Nat Genet 35 supplement no. 1, September 2003. Available at http://www.nature.com/ng/journal/v35/n1s/index.html

Human genome: organization and evolution

Bailey JA (2002) Recent segmental duplications in the human genome. *Science* 297:1003–1007; PMID 12169732.

Conrad B & Antonorakis SE (2007) Gene duplication: a drive for phenotypic diversity and cause of human disease. *Annu Rev Genomics Hum Genet* 8:17–35; PMID 17386002.

Konkel MK & Batzer MA (2010) A mobile threat to genome stability: the impact of non-LTR retrotransposons upon the human genome. *Semin Cancer Biol* 20:211–221; PMID 20307669.

Long M (2013) New gene evolution: little did we know. *Annu Rev Genet* 47:307–333; PMID 24050177.

Mills RE (2007) Which transposable elements are active in the human genome? *Trends Genet* 23:183–191; PMID 17331616.

Muotri AR (2007) The necessary junk: new functions for transposable elements. *Hum Molec Genet* 16:R159-R167; PMID 17911158.

Pink RC (2011) Pseudogenes: pseudo-functional or key regulators in health and disease. *RNA* 17:792–798; PMID 21398401.

Ponting CP & Hardison RC (2011) What fraction of the human genome is functional? *Genome Res* 21:1769–1776; PMID 21875934.

Vinckenbosch N (2006) Evolutionary fate of retroposed gene copies in the human genome. *Proc Natl Acad Sci USA* 103:3220–3225; PMID 16492757.

3

Principles underlying core DNA technologies

DOI: 10.1201/9781003044406-3

CONTENTS

Defining the genetic basis of disease requires the analysis of DNA and sometimes chromosomes. That can be challenging because the vast majority of our genetic material is organized as immensely long DNA molecules, and disease may result from a mutation that changes just a single nucleotide out of the more than six billion nucleotides in the combined maternal and paternal genomes that we inherit. Sophisticated DNA

technologies allow us to study and analyze genes, enabling diagnostic, predictive, and therapeutic applications that we describe in later chapters. Here we confine ourselves to describing the *principles* of four core technologies for purifying and analyzing DNA sequences.

Three of the four core DNA technologies began to be employed in the first efforts to study and identify individual human genes. The problem here was that human genes are often composed of very small exons separated by very large introns, and individual exons and genes represent a tiny fraction of our genome. Although we can readily isolate DNA from human cells, a single coding sequence exon, averaging just 150 bp, is a tiny fraction of the DNA (just 1/20 000 000 of the genome). Many full-length genes are also extremely small components of the genome. To allow us to focus on just a single exon or single gene, two quite different approaches can be employed. Either the DNA of interest must be purified by selectively increasing its copy number (DNA amplification), or it must be specifically recognized in some way (see **Table 3.1**).

TABLE 3.1 TWO VERY DIFFERENT APPROACHES USED TO ENABLE DETAILED STUDY OF A SHORT DNA OF INTEREST ("TARGET DNA") IN A COMPLEX GENOME

General approach	Subapproaches	Core technology
1. Purify the target DNA by selective amplification*	Selective amplification *within cells* (using a cellular DNA polymerase)	DNA cloning (Section 3.1)
	Selective amplification *in vitro* (using a purified DNA polymerase)	PCR** (Section 3.2)

* That is, selectively increase the number of copies of the target DNA.

** PCR (the polymerase chain reaction) is the most widely used way to amplify a target DNA *in vitro,* but alternative methods exist.

General approach	Subapproaches	Core technology
2. Specifically recognize the target DNA	Various (Section 3.3)	Nucleic acid hybridization (Section 3.3)

* That is, selectively increase the number of copies of the target DNA.

** PCR (the polymerase chain reaction) is the most widely used way to amplify a target DNA *in vitro,* but alternative methods exist.

Table 3.1 shows three core DNA technologies. There is of course, a fourth. DNA sequencing is the ultimate way of tracking changes in genes and DNA sequences. It used to be expensive, time-consuming, and restricted in scope. All that has changed. Now, almost two decades into the "post-genome era" (after the human genome sequence was obtained), DNA sequencing is such a hugely efficient and dominant DNA technology that we can sequence whole human genomes quickly and cheaply. Accordingly, we can now *simultaneously* analze all known genes across our genome. We consider the general principles, and the most basic of the commonly used DNA sequencing methods in Section 3.4, but we introduce the most recent DNA sequencing techniques in Chapter 11, and some medical applications in various other chapters.

3.1 AMPLIFYING DNA BY DNA CLONING

Cloning DNA in cells is a way of purifying DNA sequences and is usually carried out in bacterial cells. It can allow very many identical copies of a desired DNA sequence to be produced, enabling it to be studied or put to some use. To do that, cells are first treated to optimize the transfer of DNA molecules to be cloned into the cells, a process known as **transformation**. In each case the DNA to be cloned is first covalently joined to some *vector* DNA molecule that will help it replicate within the host cells, as detailed below. The joining of DNA fragments to vector molecules results in the formation of an artificial **recombinant DNA**. There is normally some kind

of selection or screening system that helps identify those cells that have been successfully transformed and that contain recombinant DNA.

Transformation, a key step in DNA cloning, is highly selective: when foreign DNA does get into a cell, just a *single* DNA molecule is usually taken up by a cell. A population of cells therefore serves as a sorting office that can efficiently fractionate a complex mixture of DNA fragments (Figure 3.1).

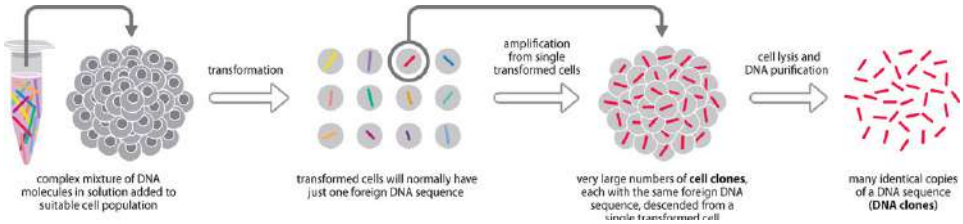


Figure 3.1 Transformation as a way of fractionating a complex sample of DNA fragments. The key point is that transformation is selective: when a cell is transformed, it usually picks up a *single* DNA molecule from the environment, and so different fragments are taken up by different cells. (For simplicity, the figure shows only the DNA sequences that are to be cloned—in practice they would be joined to a *vector DNA* molecule to make a recombinant DNA that is often circular.) Cell clones can form by cell division from a single transformed cell. Thereafter, they are propagated to produce a large number of cells with an identical foreign DNA sequence that can be purified after the cells have been broken open.

Amplifying desired DNA within bacterial cells

Large quantities of a cloned DNA can be obtained using bacterial cells because the inserted DNA can be amplified to very high copy numbers (Figure 3.1). That is possible for two reasons. First, a single bacterium containing a cloned DNA can rapidly divide and eventually produce a huge number of identical bacterial cell clones, each with the same foreign DNA sequence. Secondly, some vector molecules can replicate *within* a bacterial cell to reach quite high copy numbers; if they have a foreign DNA sequence covalently linked to them, it too will be amplified within the cell (Figure 3.2).

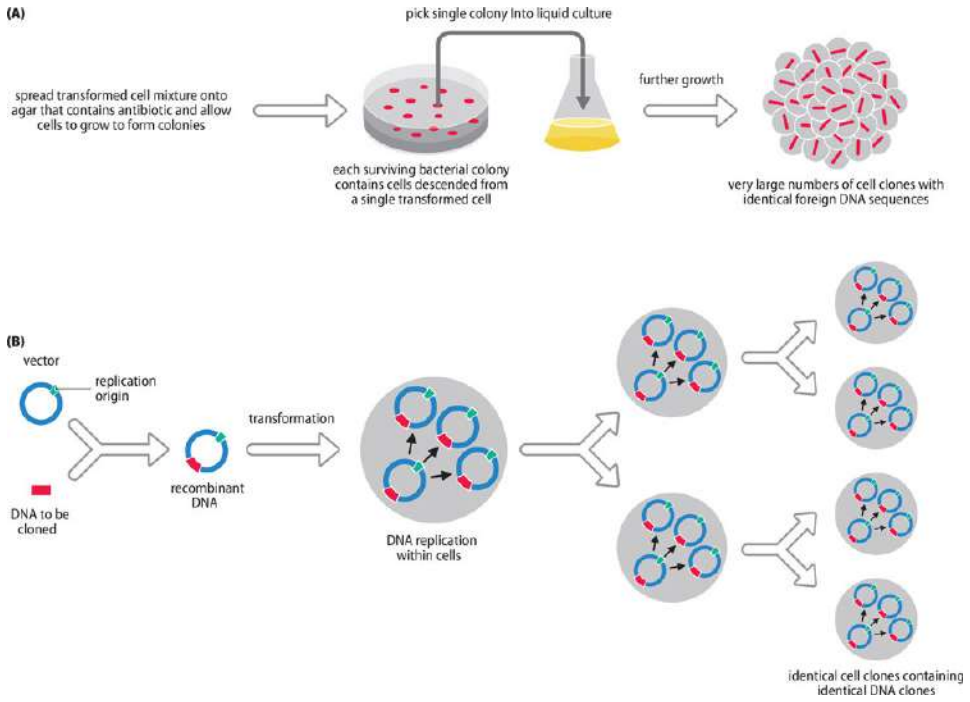


Figure 3.2 DNA cloning in bacterial cells: DNA copy number amplification and separation of clones from different transformed cells. (A) For a cell transformed by a recombinant DNA, an increase in cell number leads to a proportional expansion in copy number of that recombinant DNA. Growth occurs initially in a solid medium (after the transformed cells have been spread out on a plate of agar containing antibiotics). Individual cells will be physically dispersed on the plate, and then go through several rounds of cell division *in situ* to form *separate* visible colonies. An individual colony can be picked and allowed to go through a second round of amplification by growth in liquid culture. For simplicity, the cloned DNA fragments are shown in the absence of the vector molecule. (B) Vectors have their own replication origin allowing them to replicate *independently* of, and much more frequently than, the bacterial chromosome.

The need for vector DNA molecules

Fragments of human DNA cannot normally replicate after being transferred into bacterial cells: they lack a special DNA sequence capable of initiating DNA replication in that cell type. Such sequences are called replication origins; molecules containing a replication origin are known as ***replicons***.

To permit replication within a bacterial cell the human DNA fragment must first be covalently joined (*ligated*) to a suitable replicon, forming a

recombinant DNA. Extrachromosomal replicons are typically used for this purpose, either *plasmids* (small circular double-stranded DNAs that can replicate in bacteria) or *bacteriophages* (bacterial viruses). When used to ferry desired DNA fragments into bacterial cells, the plasmid or bacteriophage DNA is known as a **vector DNA**.

To be useful as a cloning vector the original plasmid or bacteriophage needs to be genetically modified so that we can efficiently join a foreign DNA to it (described below) and so that transformed cells can easily be recognized. The vector will often have been genetically engineered to contain a gene conferring resistance to some antibiotic to which the host bacterial cells are sensitive. After transformation, the cells are grown on agar containing the antibiotic; untrans-formed cells die, but transformed cells survive. Because some cells are transformed by naked vector DNA (lacking other DNA), screening systems are often also devised to ensure that cells with recombinant DNA can be identified.

Physical clone separation

How can cells that have taken up different DNA fragments be separated from each other? That relies on the formation of *physically separated* cell colonies. After transformation of bacterial cells, for example, aliquots of the cell mixture are spread over the surface of antibiotic-containing agar in Petri dishes (plating out); successfully transformed cells should grow and multiply; if the plating density is optimal, they form well-separated cell colonies (see Figure 3.2A). Each colony consists of identical descendant cells (cell clones) that originate from a single transformed cell and so the cell clones each contain the same single foreign DNA molecule.

An individual well-separated cell colony can then be physically picked and used to start the growth of a large culture of identical cells all containing the same foreign DNA molecule, resulting in very large amplification of a single DNA sequence of interest (Figure 3.2A). Thereafter, the cloned foreign DNA can be purified from the bacterial cells.

The need for restriction nucleases

DNA cloning in bacterial cells is most efficient when transferring relatively small DNA fragments. However, when DNA is isolated from the cells of complex organisms, the immensely long nuclear DNA molecules are fragmented by physical shearing forces to give an extremely heterogeneous collection of still rather long fragments with heterogeneous ends. The long fragments need to be reduced to pieces of a much smaller, manageable size with more uniform end sequences to facilitate ligation.

Recombinant DNA technology was first developed in the 1970s. The crucial breakthrough was to exploit the ability of restriction endonucleases to cut the DNA at *defined* places. As a result, the DNA could be reduced to small well-defined fragments with uniform end sequences that could be easily joined by a DNA ligase to similarly cut vector molecules (**Box 3.1**).

BOX 3.1 RESTRICTION ENDONUCLEASES: FROM BACTERIAL GUARDIANS TO GENETIC TOOLS

THE NATURAL ROLE OF RESTRICTION ENDONUCLEASES: HOST CELL DEFENSE

Restriction endonucleases are bacterial enzymes that recognize specific short sequence elements within a double-stranded DNA molecule, and then cleave the DNA on both strands, within, or close to, the recognition sequences. They provide a form of self-defense against bacteriophages: the restriction nuclease produced by a bacterium is designed to selectively cleave the DNA of the invading bacteriophage into small pieces, while leaving the bacterial genome intact. Different types and strains of bacteria produce restriction endonucleases of different sequence specificity.

For example, restriction nuclease *Eco*RI from the *Escherichia coli* strain RY13 specifically recognizes the sequence GAATTC and cleaves DNA strands within this recognition sequence (called a **restriction site**) – see **Figure 1A**. The same bacterial strain also initially produces an *Eco*RI methyltransferase to modify its own genome: it recognizes the same sequence GAATTC and methylates the central adenosine on both DNA

strands. The *Eco*RI restriction nuclease cannot cleave at previously methylated GAATTC sequences within the bacterial genome but will cleave at unmethylated GAATTC sequences in the DNA of invading pathogens, cutting it up into small pieces that are degraded.

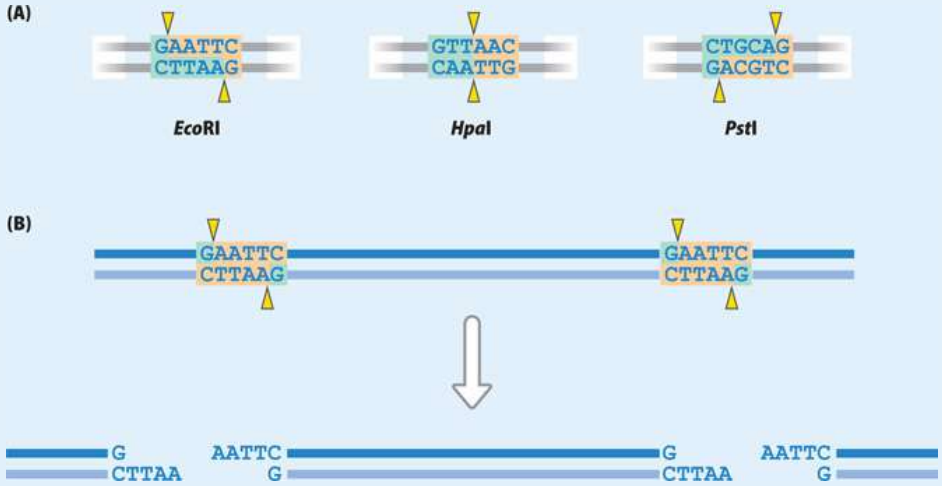


Figure 1 Sequence specificity of type II restriction nucleases. (A) Shown are the sequence specificities of three type II restriction nucleases that cleave within palindromic recognition sites. *Eco*RI (from Escherichia coli strain RY13) and *Pst*I from *Providencia stuartii* cut asymmetrically within their recognition sequences to produce four-nucleotide overhanging ends (5¢ AATT overhangs for *Eco*RI; 3¢ TGCA for *Pst*I). *Hpa*I from *Haemophilus* parainfluenzae cuts symmetrically at the middle of its recognition sequence to leave blunt-ended fragments. (B) Cleavage of genomic DNA with *Eco*RI produces fragments with two 5¢ AATT overhangs.

RESTRICTION NUCLEASES AS MOLECULAR GENETIC TOOLS

There are different classes of restriction nucleases but type II restriction nucleases are widely used in manipulating and analyzing DNA. They recognize short sequence elements that are typically *palindromes* (the 5¢ ® 3¢ sequence is the same on both strands, as in the sequence GAATTC); they then cleave the DNA either within, or very close to, the recognition sequence. Cleavage often occurs at asymmetric positions within the two strands to produce fragments with overhanging 5¢ ends or overhanging 3¢

ends, but sometimes a restriction enzyme cuts symmetrically to produce "blunt ends" (see Figure 1).

Under appropriate conditions, it is possible to use a restriction nuclease to cut complex genomic DNA into thousands or millions of fragments that can then be individually joined using a DNA ligase to a similarly cut vector molecule to produce recombinant DNA molecules (**Figure 2**). For cloning DNA in bacterial cells, vector molecules are often based on circular plasmids that have been artificially engineered so that they contain unique restriction sites for certain restriction nucleases. The recombinant DNA molecules can then be transferred into suitable host cells and amplified.

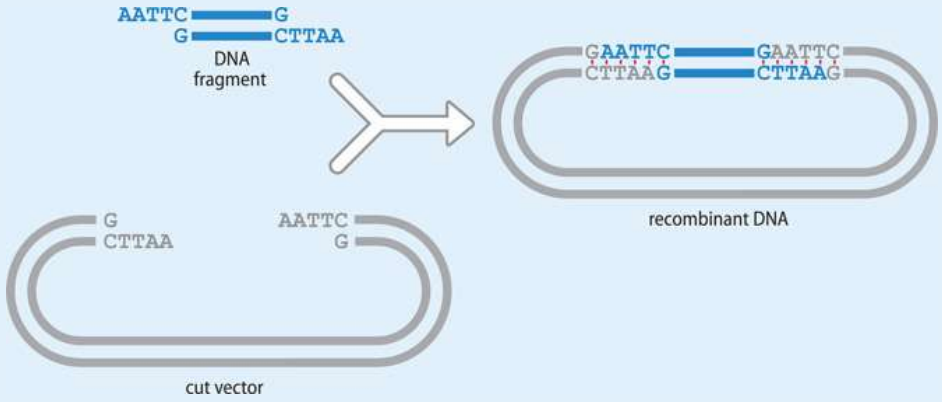


Figure 2 Formation of recombinant DNA. In this example, the vector has been cut at a unique *Eco*RI site to produce 5¢ ends with an overhanging AATT sequence, and the DNA fragment to be cloned has the same 5¢ AATT overhangs, having also been produced by cutting with *Eco*RI. The AATT overhangs are examples of *sticky ends* because they can hydrogen bond to other fragments with the same overhang thus facilitating intermolecular interactions. (Vertical red lines in the recombinant DNA represent hydrogen bonds between paired 5¢ AATT overhangs in the vector and in the DNA to be cloned.)

DNA libraries and the uses and limitations of DNA cloning

Once DNA cloning was established it was soon used to make **DNA libraries**; that is, collections of DNA clones representing all types of DNA sequence in a complex starting material.

DNA isolated from white blood cells, for example, provides a complex genomic DNA that can be cut into many pieces and attached to vector DNA molecules. The resulting mixture of different recombinant DNA molecules is used to transform bacteria to produce very many different clones, a *genomic DNA library*. A good genomic DNA library would have so many different DNA clones that there was a good chance that the library would include just about all the different DNA sequences in the genome.

An alternative was to make gene-centred *cDNA libraries* starting with mRNA. Because RNA cannot be cloned, DNA copies of the RNA were made using a specialized reverse transcriptase that naturally copies a single-stranded RNA template to make a **complementary DNA** (**cDNA**) copy. Once the cDNA strand has been made, the original RNA is destroyed by treatment with ribonuclease and the remaining DNA strand is copied in turn to give a complementary DNA, thereby making double-stranded cDNA that can be cloned like any other DNA.

DNA cloning started a revolution in genetics. It prepared the way for obtaining panels of DNA clones representing all the sequences in the genome of organisms, and that in turn made genome projects possible to obtain the complete sequence of genomic DNA in a variety of organisms. Once that was done, the structure of genes could be determined, paving the way for comprehensive studies to analyze gene expression and to determine how individual genes work.

There is a drawback: cloning DNA in cells is laborious and time-consuming. It is also not suited to performing rapid parallel amplifications of the same DNA sequence in multiple different samples of DNA. That required a new technology, as described in the next section.

3.2 AMPLIFYING DNA USING THE POLYMERASE CHAIN REACTION (PCR)

PCR, a cell-free method for amplifying DNA, was first developed in the mid-1980s and revolutionized genetics. It was both very fast and readily allowed parallel amplifications of DNA sequences from multiple starting DNA samples. If you wanted to amplify each exon of the b-globin gene from blood DNA samples from 100 different individuals with b-thalassemia, a single person could now do that in a very short time.

Basics of the polymerase chain reaction (PCR)

PCR relies on using a heat-stable DNA polymerase to synthesize copies of a small, predetermined DNA segment of interest within a complex starting DNA (such as total genomic DNA from easily accessed blood or skin cells). To initiate the synthesis of a new DNA strand, a DNA polymerase needs a single-stranded oligonucleotide **primer** that is designed to bind to a *specific* complementary sequence within the starting DNA.

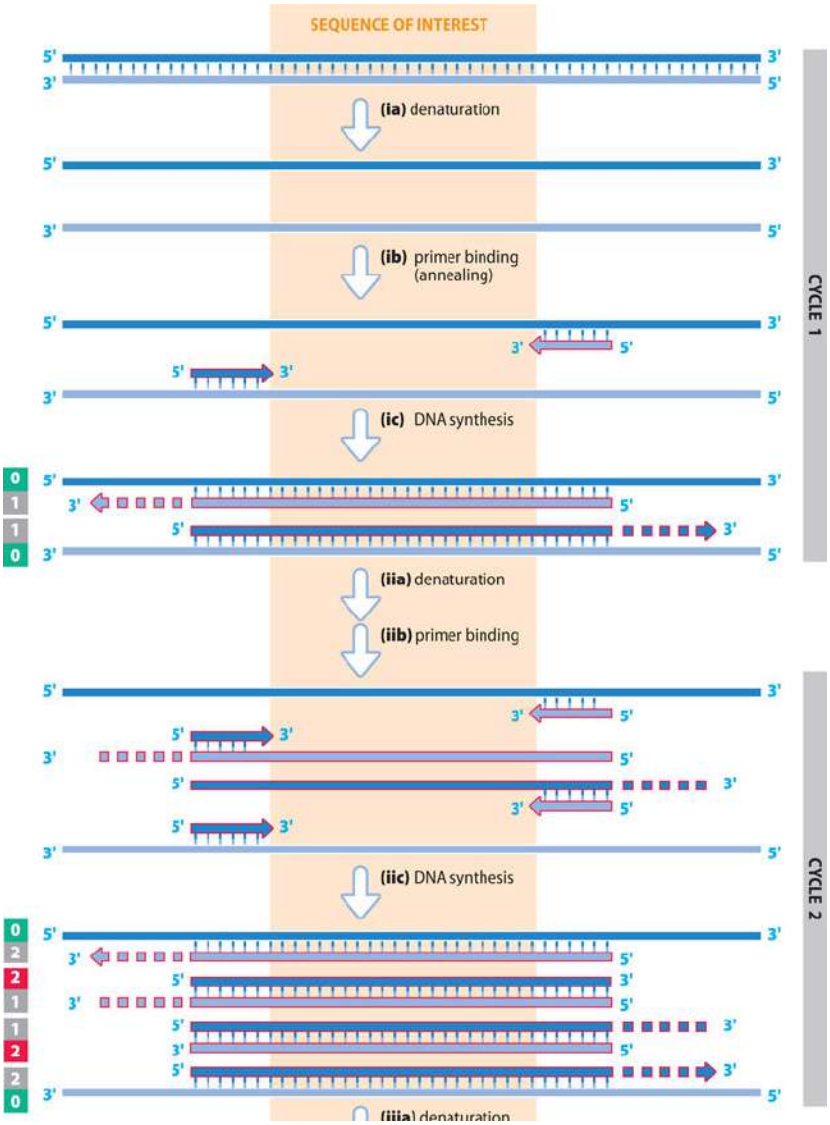
For the primer to bind preferentially at just one desired location in a complex genome, the oligonucleotide often needs to be about 20 nucleotides long or more and is designed to be able to base pair perfectly to its intended target sequence (the strength of binding depends on the number of base pairs formed and the degree of base matching).

To allow the primer to bind, the DNA needs to be heated. At a high enough temperature, the hydrogen bonds holding complementary DNA stands together are broken, causing the DNA to become single stranded. Subsequent cooling allows the oligonucleotide primer to bind to its perfect complementary sequence in the DNA sample (**annealing** or **hybridization**). Once bound, the primer can be used by a suitably heat-stable DNA polymerase to synthesize a complementary DNA strand.

In PCR, two primers are designed to bind to complementary sequences on *opposing* DNA strands, so that copies are made of both DNA strands. The primers are designed to be long enough for them to bind specifically to sequences that closely flank the DNA sequence of interest in such a way that the direction of synthesis of each new DNA stand is toward the sequence that is bound by the other primer. In further cycles of DNA

denaturation, primer binding, and DNA synthesis, the previously synthesized DNA strands become targets for binding by the other primer, causing a chain reaction to occur.

Synthesis of new DNA strands continues until the end of the template DNA is reached or until the polymerase disengages from the template DNA. The initial template DNA strands are often very long and on different copies of the template DNA the polymerase may disengage at variable places, thereby producing strands with variable 3¢ ends. However, increasingly, as the PCR reaction proceeds, template strands with fixed ends terminating in a primer sequence begin to predominate, and as a result, a product with fixed 5¢ and 3¢ ends is hugely predominant (Figure 3.3).



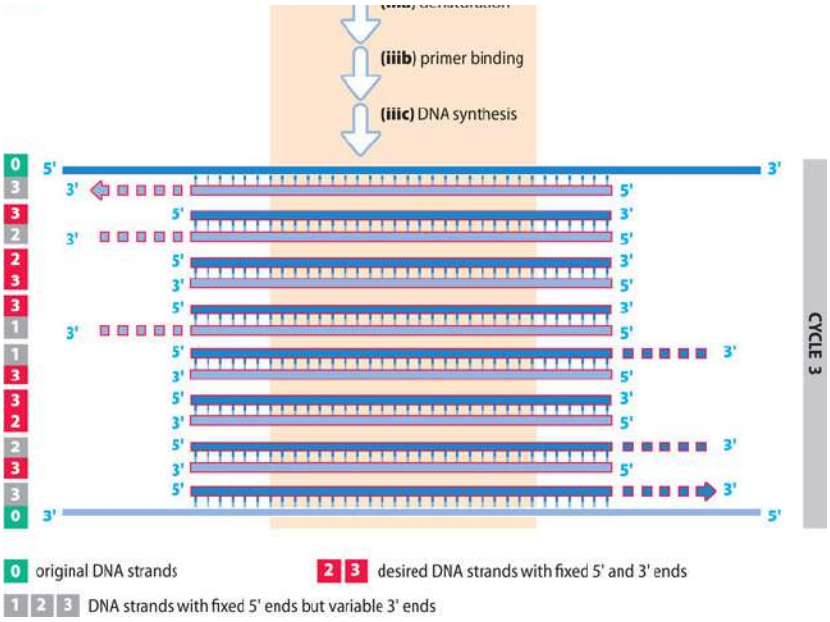


Figure 3.3 The polymerase chain reaction (PCR). The reaction usually consists of about 25–30 cycles of (a) DNA denaturation, (b) binding of oligonucleotide primers flanking the desired sequence, and (c) new DNA synthesis in which the desired DNA sequence is copied and primers are incorporated into the newly synthesized DNA strands. Numbers in the vertical strips to the left indicate the origin of the DNA strands, with original DNA strands represented by 0 and PCR products by 1 (made during first cycle), 2 (second cycle), or 3 (third cycle). The first cycle will result in new types of DNA product with a fixed 5¢ end (determined by the primer) and variable 3¢ ends (extending past the other primer). After the second cycle, there will be two more products with variable 3¢ ends but also two desired products of fixed length (shown at the left by filled red squares) with both 5¢ and 3¢ ends defined by the primer sequences. Whereas the products with variable 3¢ ends increase arithmetically (amount = *2n,* where *n* is the number of cycles), the desired products initially increase exponentially until the reaction reaches a stationary phase as the number of reactants becomes depleted (see Figure 3.4). After 25 or so cycles, the desired product accounts for the vast majority of the DNA strands.

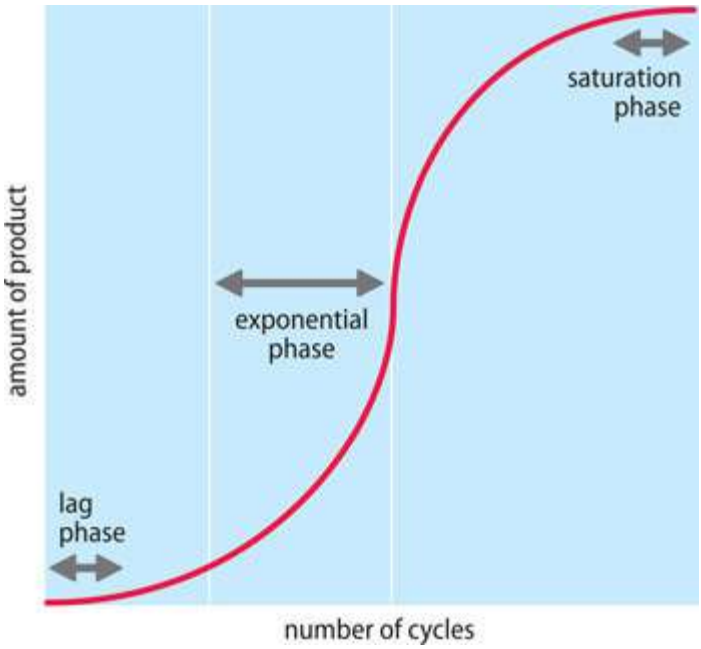


Figure 3.4 Different phases in a PCR reaction. After a lag phase, the amount of PCR product increases gradually at first. In the exponential phase, beginning after about 16–18 cycles and continuing to approximately the 25th cycle, the amount of PCR product is taken to be proportional to the amount of input DNA; quantitative PCR measurements are made on this basis. With further cycles, the amount of product increases at first but then tails off as the saturation phase approaches, when the reaction efficiency diminishes as reaction products increasingly compete with the remaining primer molecules for template DNA.

The end result is that millions of copies can be made of just the desired DNA sequence of interest within the complex starting DNA. By amplifying the desired sequence we can now study it in different ways—by directly sequencing the amplified DNA, for example.

PCR is very sensitive and can successfully amplify DNA fragments from tiny amounts of tissue samples and even from single cells. And it is robust enough to work on badly degraded tissue samples (and sometimes even samples fixed in formalin). As a result, there have been numerous applications in forensic and archaeological studies.

PCR can also be used to analyze RNA transcripts. In that case the RNA transcripts are first converted into complementary DNA (cDNA) with a

reverse transcriptase (the process is called reverse transcription-PCR or RT-PCR).

Quantitative PCR and real-time PCR

In routine PCR, all that is required is to generate a detectable or usable amount of product. However, for some purposes there is a need to quantitate the amount of product. Some **quantitative PCR** methods give a *relative quantitation* of a sequence of interest within test samples and controls, and in Chapter 11 we describe different diagnostic DNA screening methods that use PCR to get relative quantitation. Fluorescently labeled PCR products from the exponential phase of the PCR reaction (**Figure 3.4**) are removed and analyzed to measure the ratio of the fluorescence exhibited by the PCR product from a test sample (one that is associated with disease or is suspected as being abnormal) and the fluorescence exhibited by the PCR product from a control sample. The basis of the quantitation is that during the exponential phase the amount of PCR product is proportional to the amount of target DNA sequence in the input DNA.

Real-time PCR is a form of quantitative PCR that can provide absolute quantitation (the absolute number of copies), as well as relative quantitation, and is performed in specialized PCR machines. Instead of waiting for the end of the reaction, the quantitation is performed while the PCR reaction is still progressing: the amplified DNA is detected as the PCR reaction proceeds in real time within the PCR machine. Important applications are found in profiling gene expression (using RT-PCR) and also in assays for altered nucleotides in DNA, as detailed in Chapter 11.

3.3 PRINCIPLES OF NUCLEIC ACID HYBRIDIZATION

In a double-stranded DNA molecule, the hydrogen bonds between paired bases act as a fastening system that holds the two complementary DNA strands together. Two hydrogen bonds form between A and T in each A–T

base pair, and three hydrogen bonds hold G and C together in each G–C base pair (see Figure 1.3). A region of DNA that is GC-rich (having a high proportion of G–C base pairs) is therefore more stable than a region that is AT-rich.

Each hydrogen bond is individually weak, but when base matching extends over many base pairs, the cumulative strength of the hydrogen bonds becomes quite strong. (As an analogy, think of Velcro®: a single Velcro hook and loop attachment is very weak, but thousands of them make for a strong fastening system.)

Double-stranded DNA can be manipulated in different ways to break the hydrogen bonds so that the two DNA strands are separated (**denaturation**). For example, if we heat the DNA to a high enough temperature (or expose it to strong concentrations of a highly polar molecule such as formamide or urea), the hydrogen bonds break and the two complementary DNA strands separate. Subsequent gradual cooling of heat-denatured DNA allows the separated DNA strands to come together again, re-forming the base pairs in the correct order to restore the original double-stranded DNA (**Figure 3.5A**).

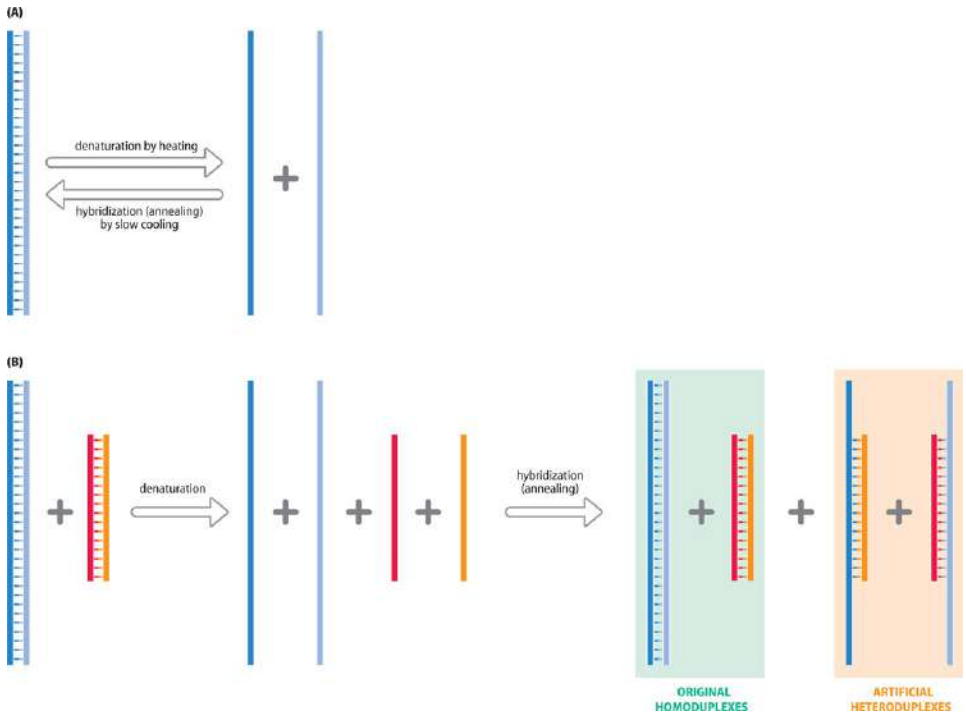


Figure 3.5 Denaturation and annealing of homologous DNA molecules to form artificial heteroduplexes and natural homoduplexes. (A) Denaturation means breaking of the hydrogen bonds in a double-stranded (duplex) nucleic acid and can be achieved by heating (or by exposure to highly polar chemicals such as urea and formamide). Under certain conditions, the separated strands can reassociate (hybridize or re-anneal) to reconstitute the original double-stranded DNA. (B) Artificial duplex formation between *homologous sequences* (in strands that have very similar nucleotide sequences) from two different DNA sources that have been denatured and mixed. For a proportion of the denatured DNA molecules, the original double-stranded DNAs re-form (homoduplexes), but in other cases artificial duplexes form between the partly complementary sequences.

Formation of artificial heteroduplexes

The association of any two complementary nucleic acid strands to form a double-stranded nucleic acid is known as nucleic acid **hybridization** (or *anneal-ing*). Under experimental conditions, two single nucleic acid strands with a high degree of base complementarity can be allowed to hybridize to form an artificial duplex. For example, if we mix cloned double-stranded DNA fragments that come from two different sources but have high levels of sequence identity, and then heat the mixture to disrupt all hydrogen bonding, all the millions of molecules of double-stranded DNA in the mixed samples from the two sources will be made single-stranded (Figure 3.5B).

Now imagine allowing the mixture to cool slowly: two different types of DNA duplex can form. First, a proportion of the single-stranded DNA molecules will base pair to their original partner to reconstitute the original DNA strands (homoduplexes). But in addition, sometimes a single-stranded DNA molecule will base pair to a complementary DNA strand in the DNA from the other source to form an artificial **heteroduplex** (see Figure 3.5B). (Note that we will use the term heteroduplex to cover all artificial duplexes in which base pairing is not perfect across the lengths of the two complementary strands. In the example in Figure 3.5B there is perfect base

matching over the length of the small DNA strands but much of the blue strands remains unpaired. Very rarely, complementary DNA strands from two different sources might be generated that have both identical lengths and perfect base matching—if so, they could form artificial homoduplexes.)

The formation of artificial duplexes, almost always heteroduplexes, is the essence of the nucleic hybridization assays that are widely used in molecular genetics. For convenience we have illustrated cloned double-stranded DNAs in Figure 3.5B. But as we will see below, the starting nucleic acids may sometimes include RNA (usually already single-stranded) or synthetic oligonucleotides as well as DNA. Often, too, one or both starting nucleic acids are complex mixtures of fragments, such as total RNA from cells or fragments of total genomic DNA. Like cloned DNA, the starting nucleic acids are usually isolated from millions of cells, and so individual sequences are normally present in many copies, often millions of copies.

Hybridization assays: using known nucleic acids to find related sequences in a test nucleic acid population

The object of a hybridization assay is to use a known nucleic acid population (the **probe**) to find related nucleic acid sequences within a poorly understood test sample. Such assays exploit the specificity of hybridization. Two single poly-nucleotide (DNA or RNA) or oligonucleotide strands will form a *stable* double-stranded hybrid (duplex) only if there is a significant amount of base pairing between them. The stability of the resulting duplex depends on the extent of base matching, and assay conditions can be chosen to allow perfectly matched duplexes only, or to allow degrees of base mismatching.

Hybridization assays can be performed in many different ways, with multiple applications in both research and diagnostics. But there is a common underlying principle: a *known,* well-characterized population of nucleic acid molecules or synthetic oligonucleotides (the *probe population*) is used to interrogate an imperfectly understood population of nucleic acids

(the test sample). To do that, both nucleic acid populations must be separated into single strands and then mixed so that single probe strands can form artificial duplexes with complementary strands in the test sample.

After the probe has bound to complementary nucleic acid strands in the test sample, the resulting probe–test-sample heteroduplexes need to be identified in some way. To do that, two conditions are needed. First, either the probe or the test-sample nucleic acid population needs to be *labeled* at the outset with modified nucleotides containing some distinctive chemical group (such as one that can emit fluorescence—we describe how nucleic acids are labeled later on in Box 3.2). **Figure 3.6** gives one approach where the probe molecules are labeled. Secondly, there must be some way of separating the probe–test-sample heteroduplexes from labeled nucleic acid homoduplexes. As described below, that usually requires that the unlabeled nucleic acids be bound to some type of solid support.

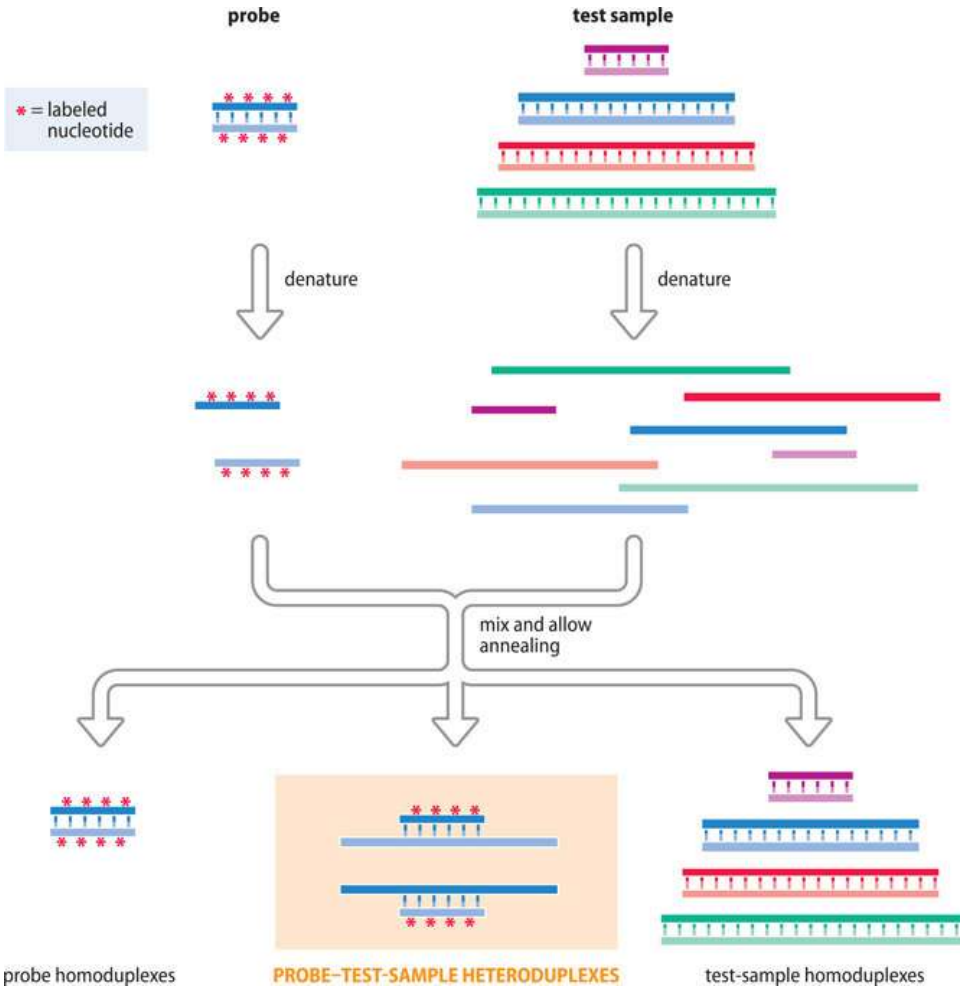


Figure 3.6 Heteroduplex formation in a nucleic acid hybridization assay. A defined probe population of known nucleic acid or oligonucleotide sequences and a test nucleic acid sample population are both made single-stranded (as required), then mixed and allowed to anneal. Many of the fragments that had previously been base paired in the two populations will reanneal to reconstitute original homoduplexes (bottom left and bottom right). In addition, new artificial duplexes will be formed between (usually) partly complementary probe and test-sample sequences (bottom center). The hybridization conditions can be adjusted to favor formation of the novel duplexes. In this way, probes can selectively bind to and identify closely related nucleic acids within a complex nucleic acid population. In this example, some kind of labeled nucleotide (*) has been introduced into the probe, but in some hybridization assays it is the test sample nucleic acid that is labeled.

Using high and low hybridization stringency

A hybridization assay can be used to identify nucleic acid sequences that are distantly related from a given nucleic acid probe. We might want to start with a DNA clone from a human gene and use that to identify the corresponding mouse gene. The human and mouse genes might be significantly different in sequence, but if we choose a long DNA probe and reduce the stringency of hybridization, stable heteroduplexes can be allowed to form even though there might be significant base mismatches (Figure 3.7A).

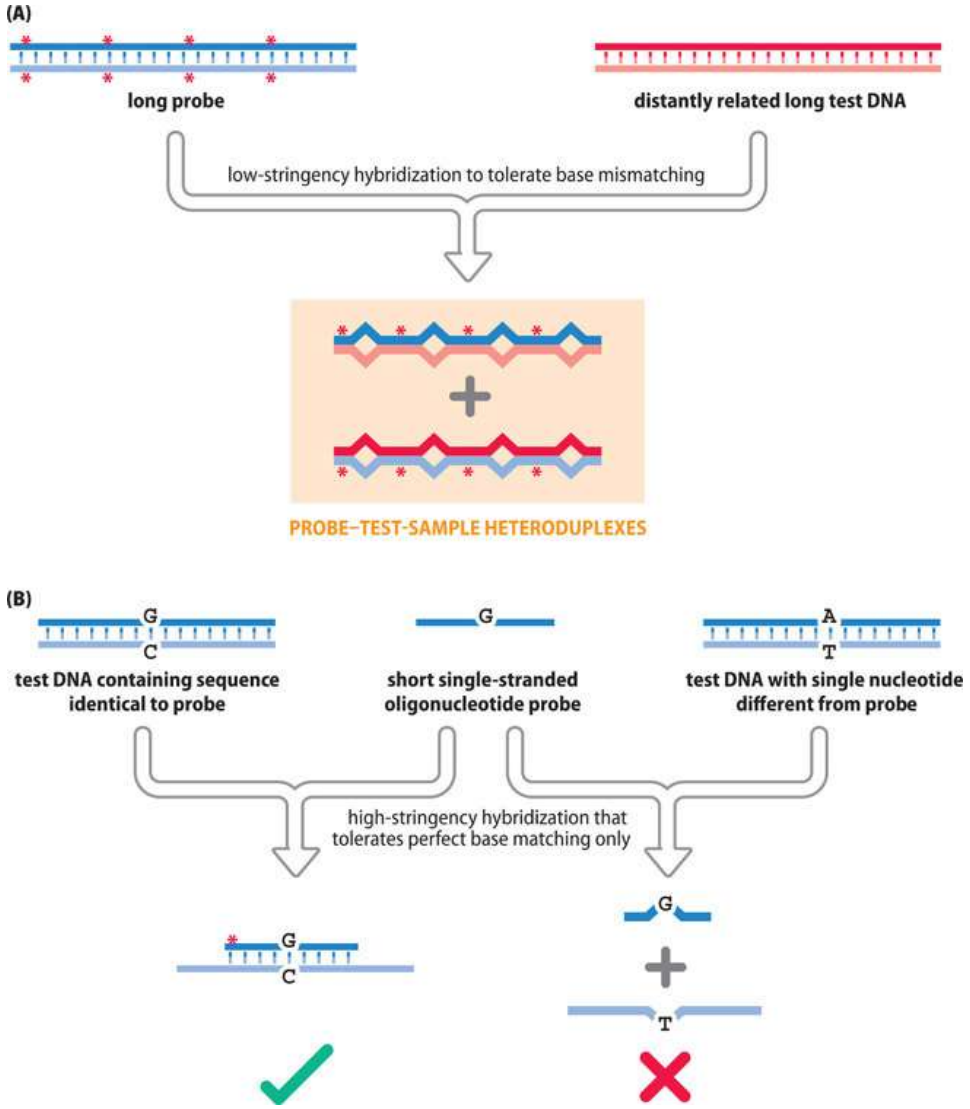


Figure 3.7 Using low or high hybridization stringency to detect nucleic acid sequences that are distantly related or show perfect base matching with a given probe. In any hybridization assay we can control the degree of base matching between complementary strands in the probe and test sample. If, for example, we increase salt concentrations and/or reduce the temperature, we lower hybridization stringency. (A) In some circumstances a long probe strand can form a thermodymically stable duplex with a comparable but distantly related strand within the test DNA (or RNA), even though there might be significant base mismatching. (B) Alternatively, we can use high temperatures and low salt concentrations to achieve high hybridization stringency that might allow only perfect base matching. That is most easily achieved with a short oligonucleotide probe and allows assays to discriminate between alleles that differ at a single nucleotide position.

Conversely, we can choose hybridization conditions to accept only perfect base matching. If we choose an oligonucleotide probe, we can use a high hybridization stringency so that the only probe-test duplexes that can form are ones that contain exactly the same sequence as the probe (Figure 3.7B). That can happen because a single mismatch out of, say, 18 base pairs can make the duplex thermodynamically unstable. Oligonucleotides can therefore be used to identify alleles that differ by a single nucleotide (*allele-specific oligonucleotides*).

Two classes of hybridization assay

There are many types of hybridization assay, but they all fall into two broad classes. In one case the probe molecules are labeled and the test-sample molecules are unlabeled, as in Figure 3.6. In that case, the probe is often a single type of cloned DNA and it is usually labeled by using a polymerase to synthesize complementary DNA or RNA strands in the presence of one or more fluorescently labeled nucleotides (**Box 3.2**). The alternative type uses unlabeled probe molecules and it is the test-sample molecules that are labeled (see below).

BOX 3.2 LABELING OF NUCLEIC ACIDS AND OLIGONUCLEOTIDES

Hybridization assays involve the labeling of either the probe or the test-sample population. Usually this involves making labeled DNA copies of a starting DNA or RNA with a suitable DNA polymerase in the presence of the four precursor deoxynucleotides (dATP, dCTP, dGTP, and dTTP). In the case of a starting RNA, a specialized DNA polymerase, a reverse transcriptase, uses the RNA as a template for making a complementary DNA copy. For some purposes, labeled RNA copies are made of a starting DNA using an RNA polymerase and the four precursor ribonucleotides (ATP, CTP, GTP, and UTP).

Whichever procedure is used, particular chemical groups (labels) are introduced into the DNA or RNA copies and can be specifically detected in some way. Often, at least one of the four nucleotide precursors has been modified so that it has a label attached to the base; alternatively, labeled oligonucleotide primers are incorporated.

Unlike DNA or RNA, oligonucleotides are chemically synthesized by the sequential addition of nucleotide residues to a starting nucleotide that will be the 3¢ terminal nucleotide. Amine or sulfhydryl groups can be incorporated into the oligonucleotide and can then be conjugated with amine-reactive or sulfhydryl-reactive labels.

Different labeling systems can be used (Table 1). Fluorescent dyes—such as derivatives of fluorescein—are popular; they can be detected readily because they emit fluorescent light of a defined wavelength when suitably stimulated. Some other labels are detected by specific binding to an antibody or to a very strongly interacting protein (see Table 1). In these cases, the detecting protein is usually conjugated to a fluorescent group (**fluorophore** or **fluorochrome**) or to an enzyme, such as alkaline phosphatase or peroxidase, which can permit detection via colorimetric assays or chemical luminescence assays.

TABLE 1 POPULAR SYSTEMS FOR LABELING NUCLEIC ACIDS

Labeling system	Examples of labels	Label detection
Fluorescence	FITC (fluorescein isothiocyanate)	using laser scanners/fluorescence microscopy
Antibody detection	digoxigenin (a steroid found in Digitalis plants)	via a digoxigenin-specific antibody that is coupled to a fluorophore or suitable enzyme
Specific protein interaction	biotin (= vitamin B7)	via streptavidin (a bacterial protein with an extraordinarily high affinity for biotin) that has been conjugated to a fluorophore or enzyme

The point of using labeled nucleic acids in a hybridization assay is to allow probe-test sample heteroduplexes to be identified. But how can we distinguish between the label in these duplexes and the label in the original labeled probe or labeled test-sample DNA? The answer is to immobilize the unlabeled nucleic acid population on a solid support (often plastic, glass, or quartz) and expose it to an aqueous solution of the labeled nucleic acid population. When labeled nucleic acid strands hybridize to complementary sequences on the solid support, they will be physically bound to the support, but labeled molecules that do not find a partner on the support or that stick nonspecifically can be washed off. That leaves behind the complementary partners that the assay is designed to find (**Figure 3.8**).

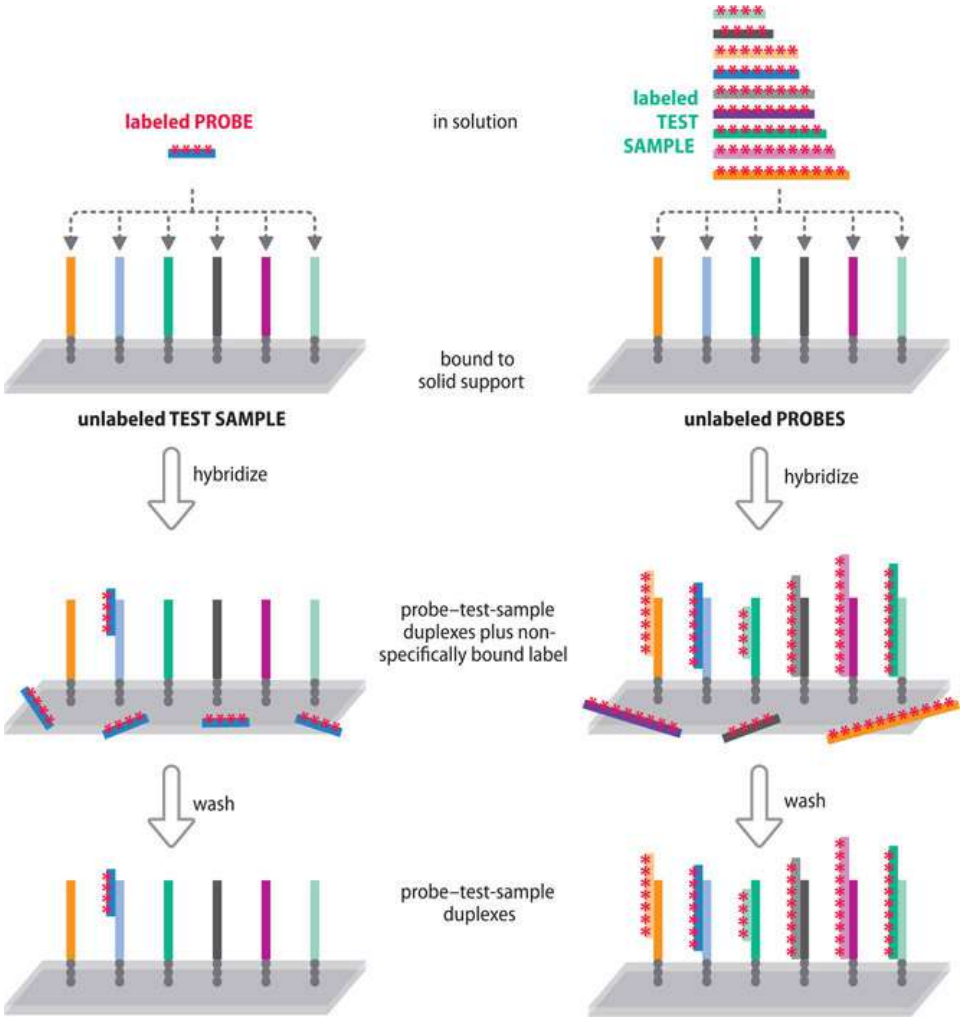


Figure 3.8 The two fundamental classes of hybridization assay and the use of solid supports to capture labeled probe-test sample duplexes. In both a standard hybridization assay, where the probe is labeled (left column), and a reverse hybridization assay, where the test sample is labeled (right column), the unlabeled nucleic acid/oligonucleotide population is bound to a solid support and denatured, before being exposed to an aqueous solution of the labeled nucleic acid/oligonucleotide population that has also, as required, been denatured. Single-stranded molecules in the labeled population can hybridize to complementary sequences in the unlabeled population, and so become bound to the solid support. Other labeled sequences that have not bound, or have bound nonspecifically at incorrect locations on the support, can be washed off. In the past, most hybridization assays were standard assays (see Table 3.2 for examples), but reverse hybridization assays became popular after microarray hybridization was developed.

TABLE 3.2 EXAMPLES OF STANDARD HYBRIDIZATION ASSAYS

Probe and test sample labeling	Hybridization method	Applications	Examples
Labeled probe and unlabeled test sample (Figures 3.6 and 3.7A)	Southern blot	looking for medium-sized changes (hundreds of base pairs to several kilobases) in genes/DNA in test sample	Clinical Box 3 Figure 2 on page 171
	tissue *in situ*	tracking RNA transcripts in tissues and embryos	
	chromosome *in situ*	studying large-scale changes using fixed chromosomes on a slide as the test sample	Figures 10.7A and 11.4

Standard hybridization assays have been used for different purposes (**Table 3.2** gives some examples). For decades, almost all hybridization assays used a homogeneous labeled probe (often, a single type of DNA clone) to search for related sequences in an immobilized complex test nucleic acid sample (see Figure 3.6 and the left part of Figure 3.8). As described in the next section, microarray-based hybridization assays use a reverse type of hybridization where unlabeled complex probe populations bound to a surface are used to interrogate a labeled test sample (the right column of Figure 3.8 shows the principle).

Microarray hybridization: large-scale parallel hybridization to immobilized probes

Innovative and powerful hybridization technologies developed in the early 1990s permit numerous hybridization assays to be conducted simultaneously on a common sample under the same conditions. A DNA or oligonucleotide microarray consists of many thousands or millions of different unlabeled DNA or oligonucleotide probe populations that have

been fixed to a glass or other suitable surface within a high-density grid format. Within each grid square are millions of identical copies of just one probe (a grid square with its probe population is called a *feature)*. For example, oligonucleotide microarrays often have a 1.28 cm × 1.28 cm surface that contain millions of different features, each occupying about 5 or 10 μm^2 (**Figure 3.9**).

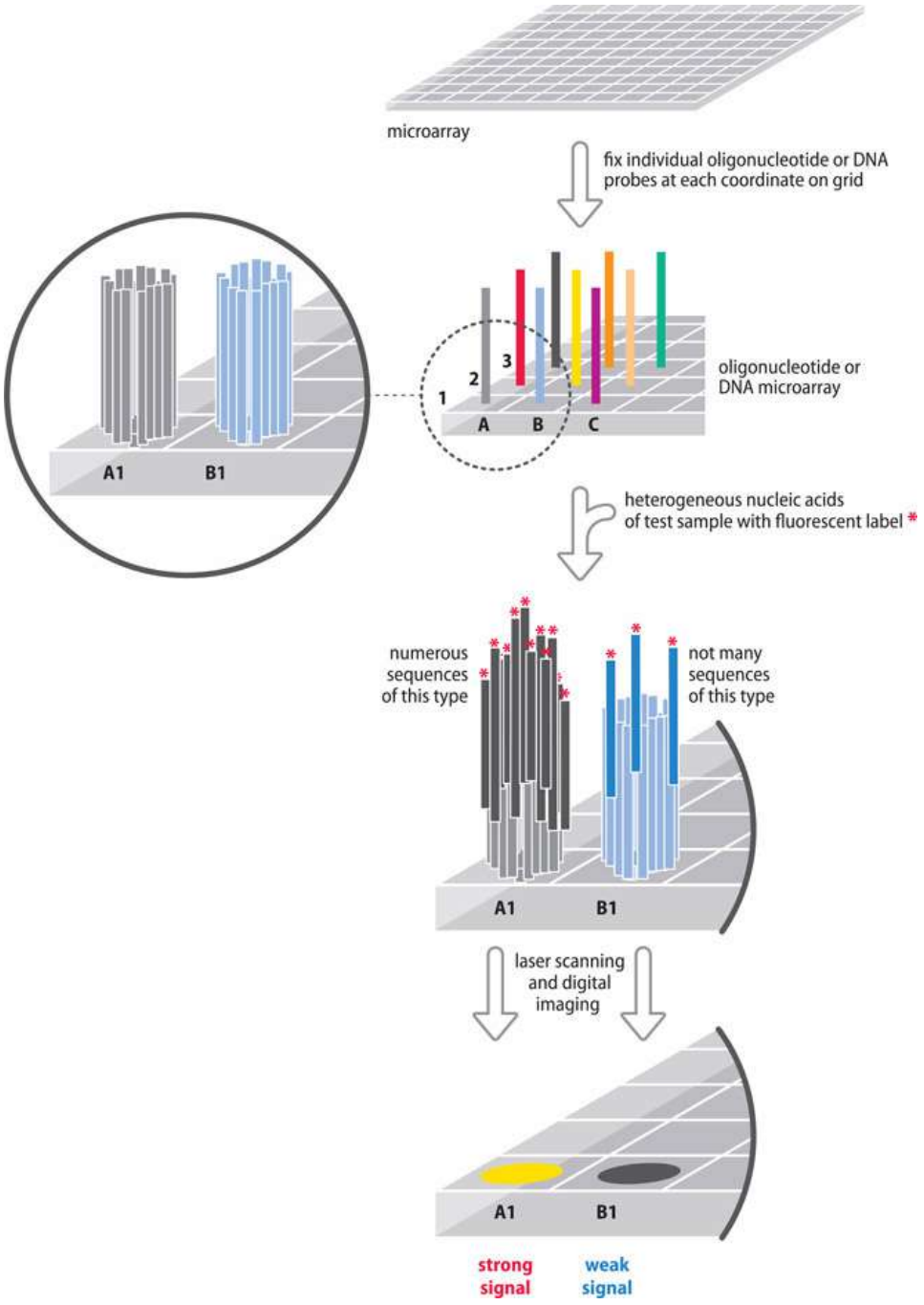


Figure 3.9 Principle of microarray hybridization. A microarray is a solid surface on which molecules can be fixed at specific coordinates in a high-density grid format. Oligonucleotide or DNA microarrays have thousands to millions of different synthetic single-stranded oligonucleotide or DNA probes fixed at specific predetermined positions in the grid. As shown by the expanded item enclosed within dashed lines, each grid square will have many thousands of identical copies of a single type of oligonucleotide or DNA probe (a *feature)*. An aqueous test sample containing a heterogeneous collection of labeled DNA fragments or RNA transcripts is denatured and allowed to hybridize with the probes on the array. Some probes (for example the A1 feature) may find numerous complementary sequences in the test population, resulting in a strong hybridization signal; for other probes (for example the B1 feature) there may be few complementary sequences in the test sample, resulting in a weak hybridization signal. After washing and drying of the grid, the hybridization signals for the numerous different probes are detected by laser scanning, giving huge amounts of data from a single experiment. (For ease of illustration, we show test-sample nucleic acids with end labels, but sometimes they contain labels on internal nucleotides.)

A test sample—an aqueous solution containing a complex population of fluorescently labeled denatured DNA or RNA—is hybridized to the different probe populations on the microarray. After a washing step to remove nonspecific binding of labeled test-sample molecules to the array, the remaining bound fluorescent label is detected with a high-resolution laser scanner. The signal emitted from each feature on the array is analyzed with digital imaging software that converts the fluorescent hybridization signal into one of a palette of colors according to its intensity (Figure 3.9).

Because the intensity of each hybridization signal reflects the number of labeled molecules that have bound to a feature, microarray hybridization is used to *quantitate* different sequences in complex test-sample populations such as different samples of genomic DNA or total cellular RNA (or cDNA). Frequent applications include quantifying different transcripts (expression profiling) and also scanning genomes to look for large-scale deletions and duplications, as described in Chapter 11.

3.4 PRINCIPLES OF DNA SEQUENCING

DNA sequencing is the ultimate DNA test. Until quite recently, Sanger dideoxy DNA sequencing was the predominant method. It relies on amplifying individual DNA sequences. For each amplified DNA, nested sets of labeled DNA copies are made and then separated according to size by gel electrophoresis.

In the last few years completely different technologies have allowed massively parallel DNA sequencing. No attempt is made to obtain the sequence of just a purified DNA component; instead, millions of DNA fragments present in a complex DNA sample are simultaneously sequenced without the need for gel electrophoresis.

Dideoxy DNA sequencing remains widely used for investigating specific DNA sequences, for example testing whether individuals have mutations in a particular gene. What the newer DNA sequencing technologies offer is a marked increase in sequencing capacity and the ability to sequence complex DNA populations, such as genomic DNA sequences, very rapidly. As a result of fast-developing technology, the running costs of DNA sequencing are plummeting, and very rapid sequencing of whole genomes is quickly becoming routine.

Dideoxy DNA sequencing

Like PCR, dideoxy DNA sequencing (also called Sanger sequencing) uses primers and a DNA polymerase to make DNA copies of specific DNA sequences of interest. To obtain enough DNA for sequencing, the DNA sequences are amplified by PCR (or sometimes by cloning in cells). The resulting purified DNAs are then sequenced, one after another, in individual reactions. Each reaction begins by denaturing a selected purified DNA. A single oligonucleotide primer is then allowed to bind and is used to make labeled DNA copies of the desired sequence (using a provided DNA polymerase and the four dNTPs).

Instead of making full-length copies of the sequence, the DNA synthesis reactions are designed to produce a population of DNA fragments sharing a common 5¢ end sequence (defined by the primer sequence) but with variable 3¢ ends. This is achieved by simultaneously having the standard dNTP precursors of DNA plus low concentrations of ddNTPs, dideoxynucleotide analogs that differ from a standard deoxynucleotide only in that they lack an OH group at the 3¢ carbon of the sugar as well as at the 2¢ carbon (**Figure 3.10A**).

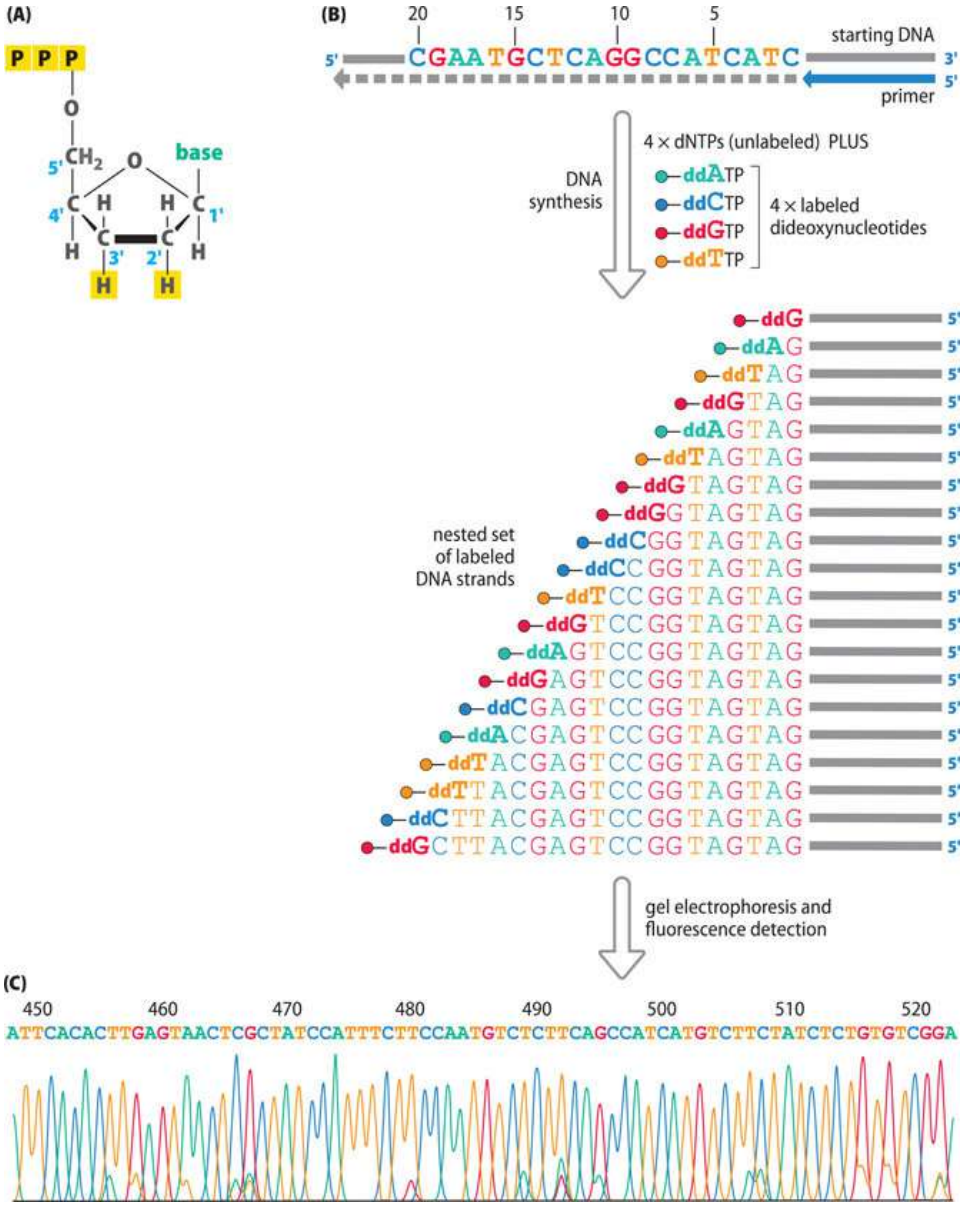


Figure 3.10 Principle of dideoxy sequencing. (A) Generalized structure of a 2¢,3¢ ddNTP. The sugar is *dideoxyribose* because the hydroxyl groups attached to both carbons 2¢ and 3¢ of ribose are each replaced by a hydrogen atom (shown by shading). (B) In dideoxy sequencing reactions, a DNA polymerase uses an oligonucleotide primer to make complementary sequences from a purified single-stranded starting DNA. The sequencing reactions include ddNTPs, which compete with the standard dNTPs to insert a chain-terminating dideoxynucleotide. Different labeling systems can be used, but it is convenient to use labeled ddNTPs that have different fluorescent groups according to the type of base, as shown here. The DNA copies will have a common 5¢ end (defined by the sequencing primer) but variable 3¢ ends, depending on where a labeled dideoxynucleotide has been inserted, producing a nested set of DNA fragments that differ by a single nucleotide in length. A series of nested fragments that differ incrementally by one nucleotide from their common 5¢ end are fractionated according to size by gel electrophoresis; the fluorescent signals are recorded and interpreted to produce a linear base sequence. (C) Example of DNA sequence output, showing a succession of dye-specific (and therefore base-specific) intensity profiles. This example shows a cDNA sequence from the *PHC3* polyhomeotic gene, provided by E. Tonkin, Newcastle University.

DNA synthesis continues smoothly when dNTPs are used, but once a dideoxynucleotide has been incorporated into a growing DNA molecule, chain extension is immediately terminated (the dideoxynucleotide lacks a 3¢-OH group to form a phosphodiester bond). To keep the balance tilted toward chain elongation, the ratio of each ddNTP to the corresponding dNTP is set to be about 1:100, so that a dideoxynucleotide is incorporated at only about 1 % of the available nucleotide positions.

If we consider competition between ddATP and dATP in the example in Figure 3.10B, there are four available positions for nucleotide insertion: opposite the T at nucleotide positions 2, 5, 13, and 16 in the starting DNA. Because the DNA synthesis reaction results in numerous DNA copies, then by chance some copies will have a dideoxyA incorporated opposite the T at position 2, some will have a dideoxyA opposite the T at position 5, and so

on. Effectively, chain elongation is *randomly* inhibited, producing sets of DNA strands that have a common 5¢ end but variable 3¢ ends.

Fluorescent dyes are used to label the DNA. One convenient way of doing this, as shown in Figure 3.10B, is to arrange matters so that the four different ddNTPs are labeled with different fluorescent dyes. The reaction products will therefore consist of DNA strands that have a labeled dideoxynucleotide at the 3¢ end carrying a distinctive fluorophore according to the type of base incorporated.

All that remains is to separate the DNA fragments according to size by electrophoresis (**Box 3.3**) and to detect the fluorescence signals. In modern dideoxy sequencing, as the DNA fragments migrate in the gel, they pass a laser that excites the fluorophores, causing them to emit fluorescence at distinct wavelengths. The fluorescence signals are recorded and an output is provided in the form of intensity profiles for the differently coloured fluorophores, as shown in Figure 3.10c.

BOX 3.3 SLAB GEL ELECTROPHORESIS AND CAPILLARY GEL ELECTROPHORESIS FOR SEPARATING NUCLEIC ACIDS ACCORDING TO SIZE

Nucleic acids carry numerous negatively charged phosphate groups and will migrate toward the positive electrode when placed in an electric field. By arranging for them to migrate through a porous gel during electrophoresis, nucleic acid molecules can be fractionated according to size. The porous gel acts as a sieve: small molecules pass easily through the pores of the gel, but larger fragments are impeded by frictional forces.

Standard gel electrophoresis with agarose gels allows the fractionation of moderately large DNA fragments (usually from about 0.1 kb to 20 kb). Pulsed-field gel electrophoresis can be used to separate much larger DNA fragments (up to megabases long). It uses specialized equipment in which the electrical polarity is regularly changed, forcing the DNA molecules to alter their conformation periodically in preparation for migrating in a

different direction. Polyacrylamide gel electrophoresis allows the superior resolution of smaller nucleic acids (it is usually used to separate fragments in size ranges up to 1 kb) and is used in dideoxy DNA sequencing to separate fragments that differ in length by just a single nucleotide.

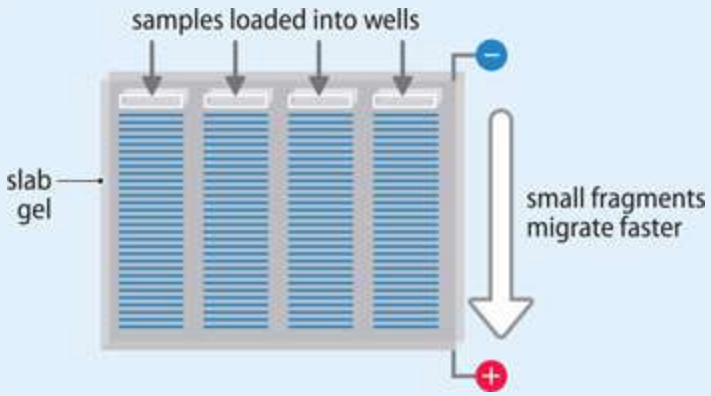


Figure 1 Slab gel electrophoresis.

In slab gel electrophoresis, individual samples are loaded into cut-out wells at one end of a solid slab of agarose or polyacrylamide gel. They migrate in parallel lanes toward the positive electrode (**Figure 1**). The separated nucleic acids can be detected in different ways. For example, after the end of an electrophoresis run, the gels can be stained with chemicals such as ethidium bromide or SYBR green that bind to nucleic acids and fluoresce when exposed to ultraviolet radiation. Sometimes the nucleic acids are labeled with fluorophores before electrophoresis, and during electrophoresis a recorder detects the fluorescence of individual labeled nucleic acid fragments as they sequentially pass a recorder placed opposite a fixed position in the gel.

The disadvantage of slab gel electrophoresis is that it is labor-intensive. The modern trend is to use capillary gel electrophoresis, which is largely automated. Fluorescently labeled DNA samples migrate through individual long and very thin tubes containing polyacrylamide gel, and a recorder detects fluorescence emissions as samples pass a fixed point (**Figure 2**). Modern dideoxy DNA sequencing uses capillary electrophoresis, as do many different types of diagnostic DNA screening methods that we outline in Chapter 11.

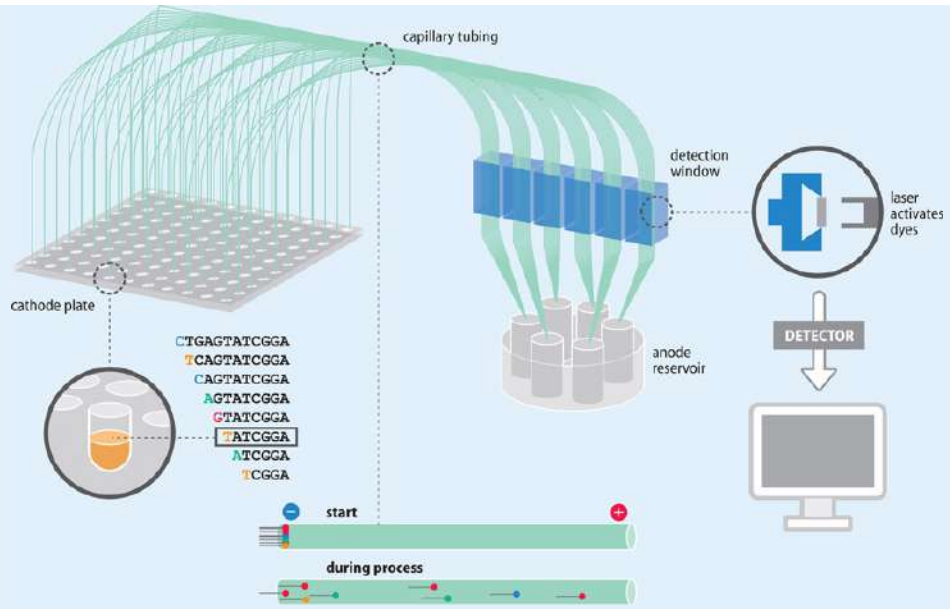


Figure 2 Capillary gel electrophoresis.

Dideoxy DNA sequencing is disadvantaged by relying on gel electrophoresis (slab polyacrylamide gels were used initially; more modern machines use capillary electrophoresis [Box 3.3]). Because gel electrophoresis is not suitable for handling large numbers of samples at a time, dideoxy sequencing has a limited sequence capacity. It is therefore not well suited to genome sequencing (although it has been used in the past to obtain the first human genome sequences). In modern times it is often used for analyzing variation over small DNA regions, such as regions encompassing individual exons.

Massively parallel DNA sequencing (next-generation sequencing)

In the early to mid-2000s new sequencing-by-synthesis methods were developed that could record the DNA sequence while the DNA strand is being synthesized. That is, the sequencing method was able to monitor the incorporation of each nucleotide in the growing DNA chain and to identify which nucleotide was being incorporated at each step.

The new sequencing technologies, often called next-generation sequencing (NGS), represent a radical step-change in sequencing technology. Standard dideoxy sequencing is a highly targeted method requiring the purification of specific sequences of interest that are then selected to be sequenced, one after another. By contrast, massively parallel DNA sequencing is indiscriminate: all of the different DNA fragments in a complex starting DNA sample can be *simul- taneously* sequenced without any need for gel electrophoresis. The difference in sequencing output is therefore vast. As listed in **Table 3.3**, various NGS technologies are commercially well established. Some of them require amplification of the starting DNA; others rely on unamplified starting DNA ("single-molecule sequencing").

TABLE 3.3 MAJOR CHARACTERISTICS OF SOME COMMERCIALLY AVAILABLE DNA SEQUENCING TECHNOLOGIES

Technology class	Sequencing platform	Read length (nucleotides)	Throughput (DNA sequence per run)
Conventional (dideoxy chain termination sequencing)	ABI prism 3730 Sanger dideoxy sequencing	~700	65 kb
Massively parallel sequencing of PCR-amplified DNAs	Illumina/Solexa NextSeq 2000	300	300 Gb
	Life Technologies Ion Torrent	200	50Gb
Massively parallel sequencing of unamplified (single-molecule) DNAs	Pacific Biosciences Sequel II	~25000	~2Gb

* Per flow cell

Technology class	Sequencing platform	Read length (nucleotides)	Throughput (DNA sequence per run)
	Oxford Nanopore MinION	> 2 000 000	30 Gb*

* Per flow cell

They vary in different parameters, such as read lengths (the length of DNA sequence generated per starting DNA), run lengths, and the number of different DNA sequences that can be conducted in parallel.

By comparison with the standard Sanger dideoxy sequencing, the NGS methods generally have high intrinsic error rates in base calling but the final reported sequences are much more accurate than the initial reads (after quality filtering and comparison of multiple sequence reads). And, importantly, they have significantly cheaper running costs per base (but are not suited to low-capacity sequencing).

A variety of additional single-molecule sequencing technologies are currently also being piloted, and sequencing capacity is likely to be increased in the near future, with yet further decreases in sequencing costs. We will describe two widely used massively parallel DNA sequencing technologies in Chapter 11.

SUMMARY

- In complex genomes, an individual gene, exon, or other sequence of interest is often a tiny fraction of the genome. To study a specific short DNA sequence like this either we must first *purify* it by selectively amplifying its copy number using some a DNA polymerase) or use some method to specifically *track* the sequence.

- Making multiple copies of a DNA sequence can be done within cells (DNA cloning), or in a cell-free system (notably by using PCR).
- In DNA cloning, the DNA sequence of interest is first attached to a vector DNA molecule that can self-replicate in a suitable host cell (often a bacterial cell). Vector molecules are modified DNAs that can readily replicate in the host cell, such as small circular plasmids or different types of bacteriophage.
- Restriction nucleases are used to cut large DNA mol ecules, such as chromosomal DNAs, into small pieces of discrete sizes that can easily be joined to similarly cut vector molecules, producing recombinant DNA molecules.
- Recombinant DNA molecules can be induced to enter a suitable host cell (transformation). Transformation is selective: each transformed cell has normally taken up a *single* DNA molecule. A transformed bacterial cell can multiply many times, and large numbers of identical copies of the recombinant DNA are produced.
- A DNA library is a bank of DNA clones that collectively include many different DNA sequences representing a complex starting population of genomic DNA (or cDNA copies of a complex RNA population).
- In PCR, a DNA sequence of interest can be copied many times from a complex source of DNA by *in vitro* DNA synthesis. Specific oligonucleotide primers are designed to bind to the starting DNA at positions flanking the sequence of interest and then used to make DNA copies that can themselves serve as templates for making further copies, rapidly increasing the copy number of the sequence of interest.
- Nucleic acid hybridization is the key method used to track a DNA or RNA sequence of interest. The method relies on the specificity of base pairing—if two different nucleic acids are

related in sequence, they may be able to form an artificial heteroduplex that is stable under selected experimental conditions.

- To perform nucleic acid hybridization, a test nucleic acid population with some sequence or sequences of interest is made single-stranded (denatured) and mixed with a probe population of known denatured nucleic acids. The object is to identify heteroduplexes in which a single-stranded sequence of interest in the test sample has formed a stable hybrid with a known sequence within the probe population.
- In many types of nucleic acid hybridization, a homo geneous labeled probe population is used to identify related sequences in an unlabeled test population that is typically bound to a solid surface.
- In microarray hybridization, many thousands of unlabeled oligonucleotide probe populations are attached to a solid surface in a regular grid formation and hybridized in parallel with a labeled test nucleic acid population provided in solution. According to the amount of labeled DNA bound to each type of oligonucleotide, it is possible to quantitate specific sequences that are complementary to each of the different probes.
- In DNA sequencing, DNA samples are made single-stranded and a DNA polymerase is used to synthesize a complementary DNA in a way that provides a readout of the base sequence.
- In standard dideoxy DNA sequencing, selected indi vidual DNA samples are sequenced. The DNA synthesis step uses a mixture of normal and chain-terminating nucleotides, producing a nested set of fragments that differ incrementally by one nucleotide and that can be separated by gel electrophoresis.

- In massively parallel DNA sequencing (next-genera tion sequencing), a complex population of very many (often millions of) DNA templates are sequenced simultaneously and indiscriminately. There is no gel electrophoresis. Instead, the methods rely on being able to monitor which of the four nucleotides is being incorporated at each step in synthesizing the cDNA.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

Brown TA (2016) *Gene cloning and DNA analyses. An Introduction*, 7th ed. Wiley-Blackwell.

Geschwind DH (2003) DNA microarrays: translation of the genome from laboratory to clinic. *Lancet Neurol* 2:275–282; PMID 12849181.

Goodwin S (2016) Coming of age: ten years of next-generation sequencing technologies. *Nature Rev Genet* 17:333–351; PMID 27184599.

4

Principles of genetic variation

DOI: 10.1201/9781003044406-4

CONTENTS

The human genome reference sequence, described in Section 2.3, is just a single representative snapshot of our genome, and an artificial one (different parts of the reference sequence originated from different people). We may speak about *the* human genome, but, in reality, there are many billions of different human genomes that owe their differences to genetic variation.

Genetic variation is very largely inherited, transmitted between generations in gametes. During life, every fertile man makes billions of sperm cells, but each sperm cell—and each egg cell—is genetically unique (pre-existing genetic variation is shuffl ed at meiosis by recombination and independent chromosome assortment). As a result, every one of us arose from a single fertilized egg cell that contained a unique diploid genome. (Occasionally, however, splitting of the very early embryo generates genetically identical embryos that can give rise to twins or, rarely, triplets.)

The vast majority of our genetic information is stored in the nuclear DNA molecules contained within our chromosomes. Each of us inherits two different haploid nuclear genomes, a paternal genome and a maternal genome, and so inherited genetic variation occurs within, as well as between, individuals. At any genetic **locus** (DNA region having a unique chromosomal location) the maternal and paternal DNA sequences (**alleles**) normally have identical or slightly different DNA sequences (we are said to be **homozygotes** if the alleles are identical, or **heterozygotes** if they differ by even a single nucleotide).

Two regions of the human genome are always inherited from a single parent. The nonrecombining portion of the Y chromosome has no sequence counterparts on the X chromosome and is transmitted exclusively by fathers to sons. Men are said to be **hemizygous** for sequences in this region, having inherited just a single allele at each locus. And all of us inherit the tiny mitochondrial DNA (mtDNA) exclusively from our mothers. (The transmitted maternal egg, however, contains about 100 000 mtDNA molecules that may show some differences in sequence; this type of mitochondrial DNA sequence variation is described as *heteroplasmy*.)

The genetic variation inherited in the fertilized egg, and present in all our nucleated cells is known as *constitutional variation*. Additional changes occur in the DNA of each of our nucleated cells throughout life, constituting *post-zygotic* or *somatic* genetic variation. (Note, however, that the qualifying term *somatic* can be used to describe genetic variation in two ways [see **Table 4.1**]).

TABLE 4.1 SOME COMMON WAYS OF CLASSIFYING GENETIC VARIATION

Classification according to:	Type	Description
Timing during development	Constitutional	Present in the zygote and transmitted to descendent cells
	Post-zygotic*	Not in the zygote, but occurring in some post-zygotic cell and transmitted to descendants of that cell
Possibility of transfer	Germline	Occurring in gametes or in any direct precursor cell.
to next generation	Somatic* (=non-germline)	Occurring in cells other than gametes or their direct precursors (such as lymphocytes, neurons, and so on)
Changes in different copies of same DNA	Allelic	Variation between maternal and paternal copies of the same chromosomal DNA sequence in a person
sequence	Heteroplasmic	Variation between different copies of the mtDNA sequence in a person

* A post-zygotic mutation is often loosely described as somatic but some occur in precursor cells in the germ line.

Most post-zygotic DNA changes occur in a rather random fashion, causing small differences in the DNA within different body cells. However, *programmed* DNA changes are also programmed to occur in certain genes in some cells, notably in maturing B and T cells (to ensure t hat each of us can make huge numbers of different antibodies and different T-cell receptors).

Individuals differ from each other mostly because our DNA sequences differ, but genetic variation is not the only explanation for differences in **phenotype** (our observable characteristics). A fertilized egg cell can split in two in early development and give rise to genetically identical twins (monozygotic twins) that nevertheless grow up to be different: although

hugely important, genetic variation is not the only influence on phenotype. During development, additional effects on the phenotype occur by a combination of stochastic (random) factors, differential gene-environment interactions and additional *epigenetic* variation that is not attributable to changes in base sequence. We consider epi-genetic effects and environmental factors in later chapters when we examine how our genes are regulated.

In this chapter we look at general principles of human genetic variation and how variation in DNA relates to variation in the sequences of proteins and non-coding RNAs. We are not concerned here with the very small fraction of genetic variation that causes disease. That will be covered in later chapters, especially in Chapter 7 (where we look primarily at genetic variation in relation to monogenic disorders), Chapter 8 (genetic variation in relation to complex inherited diseases), and Chapter 10 (genetic variation and cancer).

In Section 4.1 we consider the origins of DNA sequence variation. We take a broad look at the extent of human genetic variation in Section 4.2 and at the different forms in which this variation manifests. In Section 4.3 we deal with functional genetic variation. Here, we examine in a general way how variation in the sequences of protein products is determined both by genetic variation and by post-transcriptional modification. In this section we also deal with aspects of population genetics that relate to the spread of advantageous DNA variants through human populations (but the population genetics of harmful disease-associated DNA variants is examined in later chapters).

Genetic variation is most highly developed in genes that work in recognizing foreign, potentially harmful, molecules that have been introduced into the body. These molecules are often under independent genetic control, as in the case of infection by microbial pathogens. When that happens, two types of Darwinian natural selection may oppose each other. Thus, natural selection works on us to maximize genetic variation in the frontline immune system genes needed to recognize antigens on the invaders. Some genetic variants in these genes will be more effective than

others; accordingly, some individuals in the population will be more resistant than others to the potential harmful effects of specific microbial pathogens. But natural selection also works on the microbial pathogens to maximize genetic diversity of external molecules in an effort to escape detection by our immune defense systems.

As described in Section 4.4, the frontline genes in our immune system defenses need to recognize a potentially huge number of foreign antigens. Here, we describe the basis for the quite exceptional variability of some human leukocyte antigen (HLA) proteins and the medical significance of this variability. We also consider how exceptional post-zygotic genetic variation is created at our immunoglobulin and T-cell receptor loci so that an individual person can make a huge number of different antibodies and T-cell receptors.

4.1 DNA SEQUENCE VARIATION ORIGINS AND DNA REPAIR

Underlying genetic variation are changes in DNA sequences. **Mutation** describes both a process that produces altered DNA sequences (either a change in the base sequence or in the number of copies of a specific DNA sequence) and the outcome of that change (the altered DNA sequence). As events, mutations can occur at a wide variety of levels, and can have different consequences. They may contribute to a normal phenotype (such as height) or to a disease phenotype, and very rarely, they may have some beneficial effect. However, as explained below, the great majority of mutations have no obvious effect on the phenotype.

Mutations originate as a result of changes in our DNA that are not corrected by cellular DNA repair systems. The DNA changes are occasionally induced by exposure to certain environmental mutagens that include certain types of radiation (notably ionizing radiation and excessive ultraviolet irradiation), and also certain chemicals that we come into contact with. However, the great majority of mutations arise from endogenous

sources: both spontaneous errors in normal chromo-some and DNA function and also spontaneous chemical damage to DNA.

Mutations are inevitable. They may have adverse effects on individual organisms, causing aging and contributing to many human diseases. But they also provide the raw fuel for natural selection of beneficial adaptations that allow evolutionary innovation and, ultimately, the origin of new species.

Genetic variation arising from errors in chromosome and DNA function

Natural errors in various processes that affect chromosome and DNA function—chromosome segregation, recombination, and DNA replication—are important contributors to genetic variation. No cellular function can occur with 100 % efficiency and occasional mistakes are inevitable. Errors in the above processes may often not have harmful consequences, but some make important contributions to disease. We examine in detail how they can cause disease in Chapter 7; in this section we take a broad look into how they affect genetic variation in general.

DNA replication errors

General errors in DNA replication are unavoidable. Each time the DNA of a human diploid cell replicates, six billion nucleotides need to be inserted in the correct order to make new DNA molecules. Not surprisingly, DNA polymerases very occasionally insert the wrong nucleotide, resulting in mispaired bases (a base mismatch; the likelihood of such an error simply reflects the relative binding energies of correctly paired bases and mispaired bases).

In the great majority of cases, the errors are quickly corrected by the DNA polymerase itself. The major DNA polymerases engaged in replicating our DNA have an intrinsic 3¢ ® 5¢ exonuclease activity with a *proofreading func- tion*. If, by error, the wrong base is inserted, the

polymerases's 3¢ ® 5¢ exonuclease is activated and degrades the newly synthesized DNA strand from its 3¢ end, removing the wrongly inserted nucleotide and a short stretch before it. Then the DNA polymerase resumes synthesis again. If mispaired bases are not eliminated by the DNA polymerase, a DNA mismatch repair system is activated.

Another type of DNA replication error commonly occurs within regions of DNA where there are short tandem oligonucleotide repeats. If, for example, the DNA polymerase encounters a 30-nucleotide sequence with 15 sequential repeats of the AT dinucleotide, or 10 sequential repeats of the CAA trinucleotide, there will be an increased chance that during DNA replication a mistake is made in aligning the growing DNA strand with its template strand. A frequent result is that the template strand and newly synthesized strand pair up out of register by one (or sometimes more) repeat units, causing **replication slippage**, as detailed below. Errors like this are also often repaired successfully by the DNA *mismatch repair* system. We detail this repair system in Chapter 10 because of its importance in understanding cancer.

Although the vast majority of DNA changes caused by DNA replication errors are identified and corrected, some persist. We have many very effective DNA repair pathways, but DNA repair is also not 100 % effective: unrepaired changes in DNA sequence are an important source of mutations.

Chromosome segregation and recombination errors

Errors in chromosome segregation result in abnormal gametes, embryos, and somatic cells that have fewer or more chromosomes than normal and so have altered numbers of whole DNA molecules. Changes in chromosomal DNA copy number are not uncommon. If they occur in the germ line they often cause embryonic lethality or a congenital disorder (such as Down syndrome, which is commonly caused by an extra copy of chromosome 21), but changes in copy number of sex chromosomes are more readily

tolerated. In somatic cells, changes in chromosomal DNA copy number are a common feature of many cancers.

Various natural errors can also give rise to altered copy number of a specific sequence within a DNA strand that may range up to megabases in length. That can occur by different recombination (and recombination-like) mechanisms in which nonallelic (but often related) sequences align so that chromatids are paired with their DNA sequences locally out of register. Subsequent crossover (or sister chromatid exchange) produces chromatids with fewer or more copies of the sequences. The ensuing duplication or deletion of sequences may, or may not, have functional consequences—we cover the mechanisms and how they can result in disease in Chapter 7.

Various endogenous and exogenous sources can cause damage to DNA by altering its chemical structure

DNA is a comparatively stable molecule. Nevertheless, there are constant threats to its integrity, causing breakage of covalent bonds within DNA or inappropriate bonding of chemicals to DNA. Most of the damage originates spontaneously within cells (normal cellular metabolism generates some chemicals that are harmful to cells). A minority of the damage is induced by external sources.

Chemical damage to DNA can involve the cleavage of covalent bonds in the sugar-phosphate backbone of DNA, causing single-strand or double-strand breaks. Alternatively, bases are deleted (by cleavage of the N-glycosidic bond connecting a base to a sugar) or they are chemically modified in some way: certain chemical groups on bases may be replaced, or chemical groups may be added to bases, or unusual covalent bonds may form between two bases on the same strand or on complementary strands (*cross-linking*)—see **Figure 4.1** for examples. The chemically modified bases may block DNA or RNA polymerases, and cause base mispairing; if not repaired, they may induce mutations.

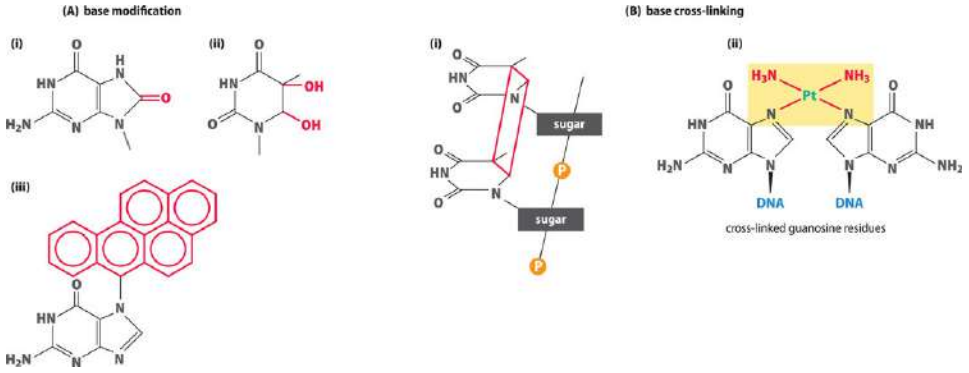


Figure 4.1 Examples of base modification and cross-linking. (A) Base modification. Altered bonding or added chemical groups are shown in red. Examples are: 8-oxoguanine (i), which base pairs to adenine and so induces mutations; thymidine glycol (ii), which is not a mutagen but blocks DNA polymerase; and a DNA adduct (iii) formed by covalent bonding, in this case of an aromatic hydrocarbon such as benzo(*a*) pyrene to N7 of a guanine residue. (B) Base cross-linking. Cyclobutane pyrimidine dimers, the most prevalent form of damage induced by solar UV light, arise by bonding of carbon atoms 4 and 5 on adjacent pyrimidines on a DNA strand (i). The anticancer agent cisplatin, (NH3)2PtCl2, causes interstrand cross-links by covalently bonding the N7 nitrogen atoms of guanines on opposite strands (ii).

Spontaneous and environmentally induced DNA damage

Most of the chemical damage to our DNA arises spontaneously and is unavoidable. Every day, under normal conditions, around 20 000–100 000 lesions are generated in the DNA of each of our nucleated cells. Hydrolytic damage can disrupt bonds that hold bases to sugars to produce an abasic site (depurination is particularly common), and also strips amino groups from some bases (deamination). Oxidative damage is also very common because normal cellular metabolism generates some strongly electrophilic (and therefore highly reactive) molecules or ions. The most significant are **reactive oxygen species** (**ROS**) formed by the incomplete one-electron reduction of oxygen, including superoxide anions ($O^{-}2$), hydrogen peroxide ($H2O2$), and hydroxyl radicals ($OH\cdot$). ROS are generated in different cellular locations and have important natural roles in certain inter-cellular

and intracellular signaling pathways, but they mostly originate in mitochondria (where electrons can prematurely reduce oxygen).

A minority of the chemical damage to our DNA is caused by external **mutagens,** agents that can induce mutation, including radiation and harmful chemicals. Ionizing radiation (X-rays, gamma rays, and so on) interacts with cellular molecules to generate ROS that break DNA strands (see below). Non-ionizing ultraviolet radiation causes covalent bonding between adjacent pyrimidines on a DNA strand (see Figure 4.1Bi).

Our bodies are also exposed to harmful environmental chemicals in our food and drink, in the air that we breathe, and so on. Some chemicals interact with cellular molecules to generate ROS. Other chemicals covalently bond to DNA, often forming bulky DNA adducts that distort the double helix. Large aromatic hydrocarbons found in cigarette smoke and automobile fumes can bond to DNA (see the example in Figure 4.1Aiii). Electrophilic alkylating agents can result in base cross-linking.

The wide range of DNA repair mechanisms

Cells have different systems for detecting and repairing DNA damage, according to the type of DNA damage. Some types of DNA damage may be minor—the net effect might simply be an altered base. Others, such as DNA cross-linking, are more problematic: they may block DNA replication (the replication fork stalls) or block transcription (the RNA polymerase stalls).

Different molecular sensors identify different types of DNA damage, triggering an appropriate DNA repair pathway. If the DNA lesion is substantial and initial repair ineffective, cell cycle arrest may be triggered that may be temporary (we consider this in the context of cancer in Chapter 10), or be more permanent. In other cases, as often happens in lymphocytes, apoptosis may be triggered.

The DNA repair process very occasionally involves a simple reversal of the molecular steps causing DNA damage. Usually, however, DNA repair pathways do not directly reverse the damage process. Instead, according to the type of lesion, one of several alternative DNA repair pathways is used.

Most of the time, the repair needs to be made to one DNA strand only; sometimes both strands need to be repaired, as in interstrand cross-linking and double-strand DNA breaks.

Errors in DNA replication and chemical damage to DNA are a constant throughout life. Inevitably, however, some mistakes are made in repairing DNA, and there are also inherent weaknesses in detecting some base changes, as described below. Inefficiency in detecting and repairing DNA damage is an important contributor to generating mutation. We consider the health consequences of defective DNA repair in **Clinical Box 1**, at the end of this section. Before this, we consider the major DNA repair mechanisms in the next two subsections.

Repair of DNA damage or altered sequence on a single DNA strand

DNA damage or an error in DNA replication usually results in one strand having a DNA lesion or a wrongly inserted base but leaves the complementary DNA strand unaffected at that location. In that case, the undamaged complementary strand may be used as a template to direct accurate repair.

- *Base excision repair* (*BER*). This pathway is specifically aimed at lesions where a single base has either been modified or excised by hydrolysis to leave an abasic site).
- *Single-strand break repair*. Simple single-strand breaks—also called *DNA nicks*—are caused by breakage of a single phosphodiester bond and are common. They are easily reversed by DNA ligase. More complex breaks occur when oxidative attack causes deoxyribose residues to disintegrate. A type of base excision repair is then employed, whereby strand breaks are rapidly detected and briefly bound by a sensor molecule, poly(ADP-ribose), that initiates repair by attracting suitable repair proteins to the site.
- *Nucleotide excision repair* (*NER*). This mechanism allows the repair of bulky, helix-distorting DNA lesions. After the lesion is detected,

the damaged site is opened out and the DNA is cleaved some distance away on either side of the lesion, generating an oligonucleotide of about 30 nucleotides containing the damaged site, which is discarded. Resynthesis of DNA is performed with the opposite strand as a template. The priority is to rapidly repair bulky lesions that block actively transcribed regions of DNA. A specialized subpathway, *transcription-coupled repair*, initiates this type of repair after detection of RNA polymerases that have stalled at the damaged site. Otherwise, an alternative global genome NER pathway is used.

- *Base mismatch repair*. This mechanism corrects errors in DNA replication. Errors in base mismatch repair are important in cancer and we describe this mechanism in Chapter 10.

Repair of DNA lesions that affect both DNA strands

Double-strand DNA breaks (DSBs) are normally rare in cells. They can occur by accident, as a result of chemical attack on DNA by endogenous or externally induced reactive oxygen species (but at much lower frequencies than SSBs). DNA repair is required but can sometimes be difficult to perform: when the two complementary DNA strands are broken simultaneously at sites sufficiently close to each other, neither base pairing nor chromatin structure may be sufficient to hold the two broken ends opposite each other. The DNA termini will often have sustained base damage and the two broken ends are liable to become physically dissociated from each other, making alignment difficult.

Unrepaired DSBs are highly dangerous to cells. The break can lead to the inactivation of a critically important gene, and the broken ends are liable to recombine with other DNA molecules, causing chromosome rearrangements that may be harmful or lethal to the cell. Cells respond to DSBs in different ways. Two major DNA repair mechanisms can be deployed to repair a DSB, as listed below; if repair is incomplete, however, apoptosis is likely to be triggered.

- *Homologous recombination* (*HR*)*-mediated DNA repair*. This highly accurate repair mechanism requires a homologous intact DNA strand to be available to act as a template strand. Normally, therefore, it operates after DNA replication (and before mitosis), using a DNA strand from the undamaged sister chromatid as a template to guide repair (Figure 4.2). It is important in early embryogenesis (when many cells are proliferating rapidly), and in the repair of proliferating cells after the DNA has replicated.

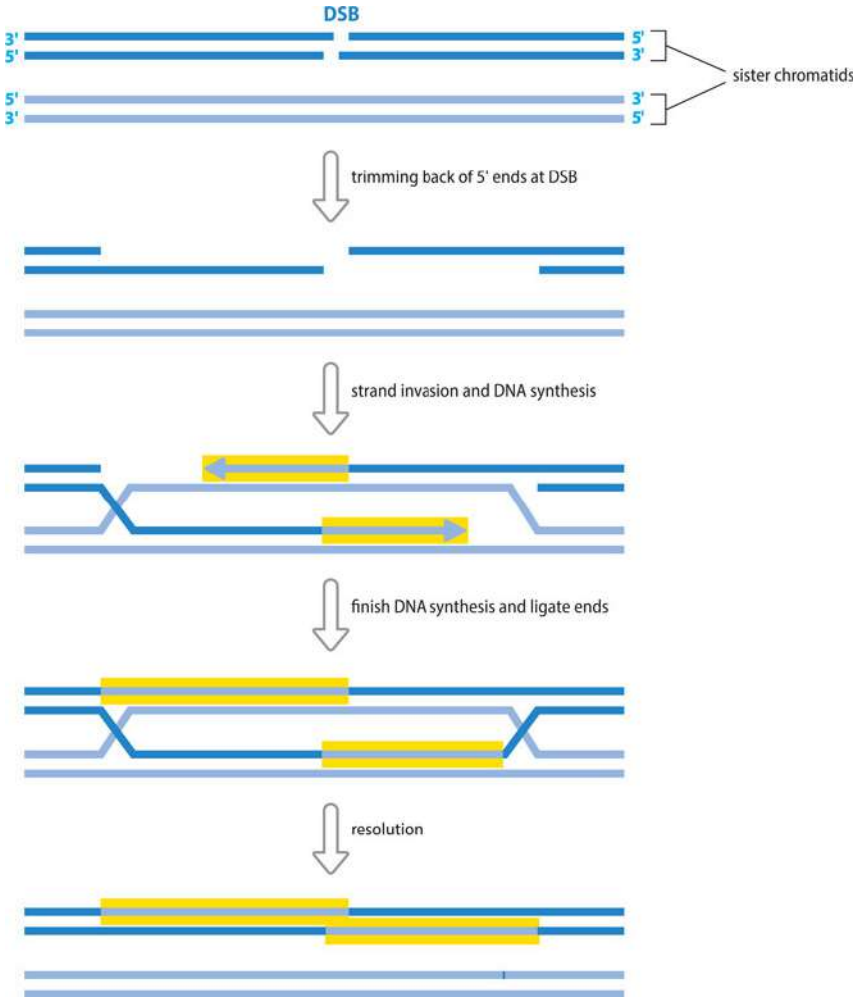


Figure 4.2 Homologous recombination-mediated repair of double-strand DNA breaks. The double-strand break (DSB) in the chromatid at the top is repaired using as a template the undamaged DNA strands in the sister chromatid (*note: to make the mechanism easier to represent, the upper chromatid has, unconventionally, the 3′ →5′ strand placed above the 5′ → 3′*

strand). The 5¢ ends at the DSB are cut back to leave protruding single-strand regions with 3¢ ends. After strand invasion, each single-strand region forms a duplex with an undamaged complementary DNA strand from the sister chromatid, which acts as a template for new DNA synthesis (shown by the arrows highlighted in yellow). After DNA synthesis, the ends are sealed by DNA ligase (newly synthesized DNA copied from the sister chromatid DNA is highlighted in yellow). The repair is highly accurate because for both broken DNA strands the undamaged sister chromatid DNA strands act as templates to direct the incorporation of the correct nucleotides during DNA synthesis.

- **Nonhomologous end joining (NHEJ).** No template strand is needed here because the broken ends are fused together. Specific proteins bind to the exposed DNA ends and recruit a special DNA ligase, DNA ligase IV, to rejoin the broken ends. Unlike HR-mediated DNA repair, NHEJ is, in principle, always available to cells. However, it is most important for the repair of differentiated cells and of proliferating cells in G1 phase, before the DNA has replicated.

Undetected DNA damage, DNA damage tolerance, and translesion synthesis

DNA damage may sometimes go undetected. In vertebrates cytosines occurring within the dinucleotide CG are highly mutable due to inefficient DNA repair. The CG dinucleotide is a frequent target for DNA methylation, converting cytosine to 5-methylcytosine (5-meC). Deamination of cytosine residues normally produces uracil, which is efficiently recognized as a foreign base in DNA and eliminated by uracil DNA glycosylase. Deamination of 5-meC, however, produces a normal DNA base, thymine, that may go undetected as an altered base (**Figure 4.3**). As a result, C ® T substitutions are the most frequent type of single-nucleotide change in our DNA.

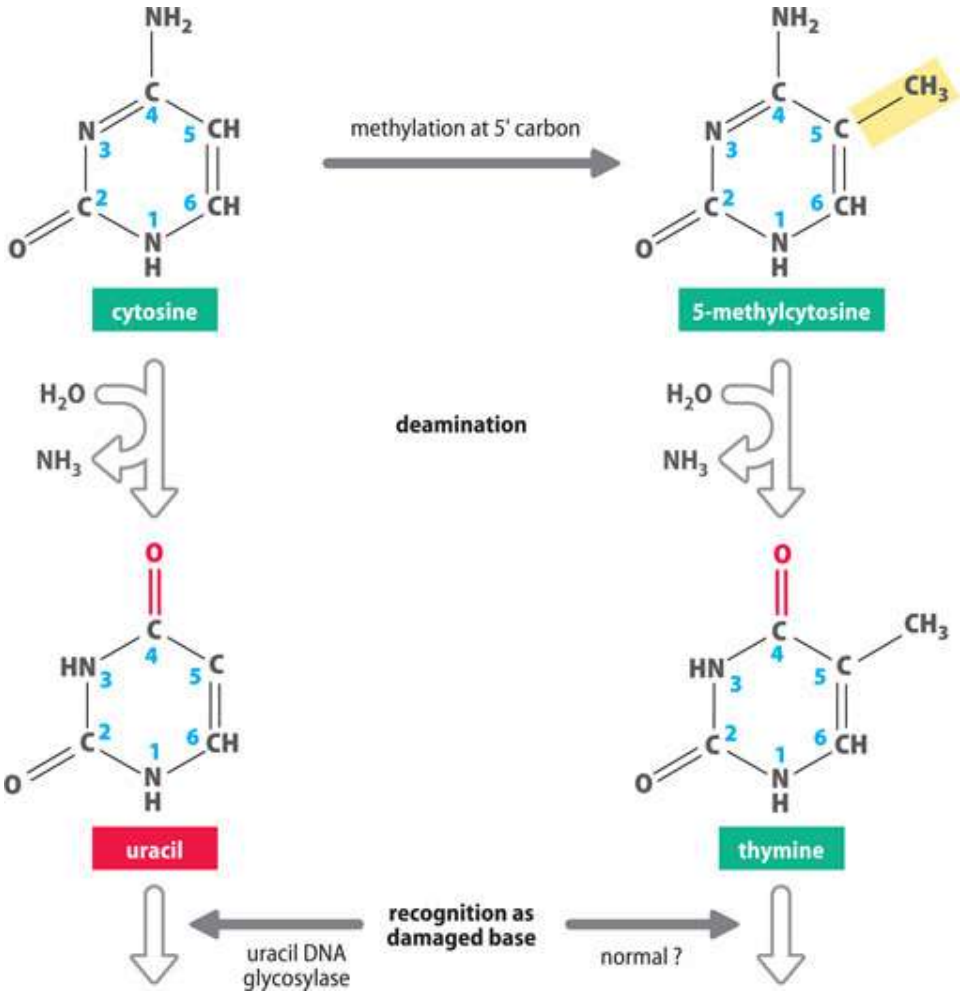


Figure 4.3 Why C → T mutations are so common in human and vertebrate DNA. Deamination of cytosine is a very common reaction in our cells and normally produces uracil, a base usually found in RNA, not DNA. In our cells a specialized enzyme, uracil DNA glycosylase, recognizes uracil residues in DNA and removes them. However, as in the DNA of other vertebrates, many of our cytosines are methylated at carbon atom 5. Deamination of 5-methylcytosine produces thymine, a base normally found in DNA. Although a stable CG base pair has been replaced by a TG base mismatch, the base mismatch may often escape detection by the base mismatch repair system (which focuses on DNA replication events). At the subsequent round of DNA replication the thymine will form a TA base pair, effectively producing a C ® T mutation.

Sometimes, DNA lesions may be identified but are not repaired before DNA replication (damage tolerance). For example, DNA lesions that block replication may be bypassed rather than repaired, and non-classical DNA

polymerases are required to resume DNA synthesis past the damaged site (*translesion synthesis*). Subsequently, the gap in the daughter strand opposite the lesion is filled in; the lesion can be repaired later on, by using the daughter strand as a template in nucleotide excision repair. The nonclassical DNA polymerases used in translesion synthesis exhibit a low fidelity in DNA replication They have a higher success in incorporating bases opposite a damaged site, but they are prone to error by occasionally inserting the wrong base. As a result, replication forks are preserved, but at the cost of mutagenesis.

CLINICAL BOX 1 DISEASES ARISING FROM DEFECTIVE DNA DAMAGE RESPONSE/DNA REPAIR

DNA damage accumulates in all of us throughout our lives. Inevitably, as we grow older, the incidence of somatic mutations increases, with consequences for increased risk of developing cancer and of declining efficiency in a variety of cellular processes, contributing to the aging process. More than 170 human genes are known to be involved in DNA damage responses and DNA repair (see Further Reading), and a wide variety of single-gene disorders are known to result from germline mutations in genes that work in these pathways (Table 1 gives some examples).

TABLE 1 EXAMPLES OF INHERITED DISORDERS OF DNA REPAIR/DNA DAMAGE RESPONSES

DNA repair/DNA damage response system	Associated single-gene disorders	Disease features*			
		C	P	N	I

* C, cancer susceptibility; P, progeria; N, neurological features; 1, immunodeficiency. HNPC, hereditary nonpolyposis cancer. SCA, spinocerebellar ataxia

DNA repair/DNA damage response system	Associated single-gene disorders	Disease features*			
		C	P	N	I
Mismatch repair (described in Chapter 10)	HNPC (Lynch syndrome)	+	-	-	-
Nucleotide excision repair (NER)	xeroderma pigmentosum	+	-	+	-
NER (transcription-coupled repair)	Cockayne syndrome	-	+	+	-
	trichothiodystrophy	-	+	+	-
Single-strand break (SSB) repair	ataxia oculomotor apraxia 1	-	-	+	-
	spinocerebellar ataxia with axonal neuropathy 1	-	-	+	-
Interstrand cross-link repair	Fanconi anemia	+	+	+	+
Double-strand break (DSB) repair (NHEJ)	Lig4 syndrome	+	-	+	+
	severe combined immunodeficiency	-	-	-	+
DNA damage signaling/DSB repair	ataxia telangiectasia	+	-	+	+
	Seckel syndrome	-	-	+	+
	primary microcephaly 1	-	-	+	-
Homologous recombination (HR)	Bloom syndrome	+	-	+	+
Base excision repair (BER) in mtDNA	spinocerebellar ataxia-epilepsy	-	-	+	-
	progressive external opthalmoplegia	-	-	-	-
Telomere maintenance (TM)	Dyskeratosis congenita	+	+	+	+
HR, BER,TM	Werner syndrome	+	+	-	-

* C, cancer susceptibility; P, progeria; N, neurological features; 1, immunodeficiency. HNPC, hereditary nonpolyposis cancer. SCA, spinocerebellar ataxia

As expected, increased susceptibility to cancer and accelerated aging are often found in these disorders, and developmental abnormalities and neurological features are also common. Although many cell types are regularly replaced, nondividing neurons are especially vulnerable. They have high oxygen and energy needs (with a resulting high frequency of oxidative damage), and they accumulate DNA damage over very long periods. Cellular abnormalities are frequently seen, with respect to chromosome and genome instability as listed below.

DISEASE FEATURES

Cancer (C) *susceptibility*. This is apparent in many inherited DNA repair deficiencies. Genome instability in mismatch repair deficiencies can induce cancer in highly proliferating tissues, notably intestinal epithelium. Individuals with xeroderma pigmentosum have little protection against UV radiation, and exposure to sunlight induces skin tumors (Figure 1A).

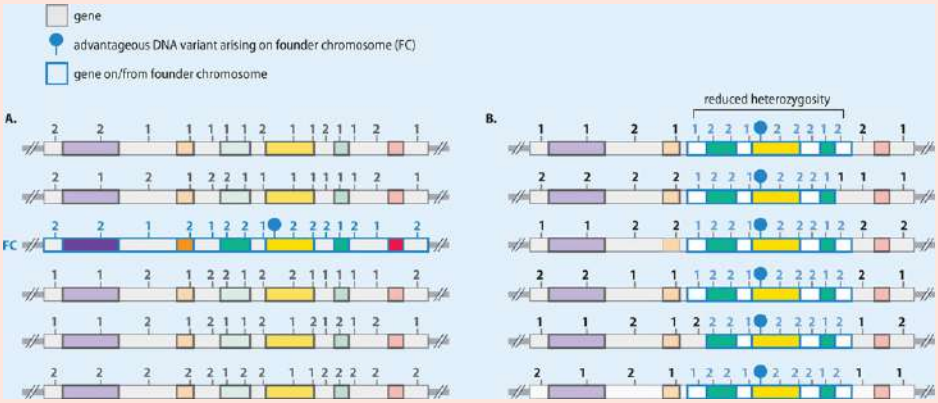


Figure 1 General effect of a selective sweep for an advantageous DNA variant. (A) Heterozygosity profile before selection. Imagine that an advantageous DNA variant has just occurred in the gene shown in yellow on a founder chromosome 22 (FC) (with genes outlined in blue). Now imagine assaying genetic variation by using intronic and extragenic microsatellite markers, each with four common alleles (1 to 4), over each copy of chromosome 22 in the population. We might expect significant heterozygosity, as shown by the six representative chromosome 22s. (B) Heterozygosity profile after positive selection over many generations. Vertical transmission of the founder chromosome 22, recombination, and continued positive

selection for the advantageous variant will result in an increased frequency of the advantageous DNA variant plus closely linked DNA variants, causing reduced heterozygosity for that chromosome segment. Some tightly linked neighboring genes will also increase in frequency in the population because of selection for the variant. They are often described as *hitchhiking alleles* (shown here in blue and green).

- *Progeria* (**P**). Some disorders have clinical features that mimic accelerated aging, notably individuals with Werner syndrome (Figure 1B), who prematurely develop gray hair, cataracts, osteoporosis, type 2 diabetes, and atherosclerosis, and generally die before the age of 50 as a result of cancer or atherosclerosis.
- *Neurological* (**N**) *features*. Neuronal death and neuro-degeneration are common features. Individuals with ataxia telangiectasia experience cerebellar degeneration leading to profound ataxia and become confined to a wheelchair before the age of 10. Microcephaly is common, sometimes accompanied by neurodegeneration and learning difficulties.
- *Immunodeficiency* (**I**). Some DNA repair proteins also work in specialized genetic mechanisms in B and T lymphocytes. For example, components of the NHEJ repair pathway are needed to make immunoglobulin and T-cell receptors, and when they are lacking, hypogammaglobulinemia and lymphopenia or severe combined immunodeficiency result.

CELL ANALYSES REVEALING GENOME AND CHROMOSOMAL INSTABILITY

The DNA of individuals with disorders of mismatch repair (described in Section 10.3) shows striking evidence of genome instability when short tandem repeat DNA polymorphisms known as microsatellite DNA are assayed. Cells from individuals with a DNA repair disorder quite often also show an increased frequency of spontaneous chromosome aberrations characteristic of the disorder, as in the case of ataxia telangiectasia,

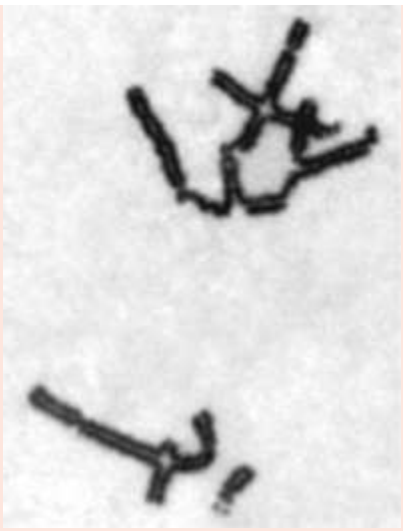
Fanconi anemia, and Bloom syndrome (which shows very high levels of sister chromatid exchange).

Chromosome analyses can also provide a simple route to laboratory-based diagnosis. Fanconi anemia (where there are variably assorted developmental abnormalities, plus progressive bone marrow failure and an increased risk of malignancy) can be caused by mutations in any one of at least 13 different genes that repair interstrand cross-links, making DNA-based diagnosis difficult. Chromosome-based diagnosis is more straightforward: lymphocyte cultures are treated with diepoxybutane or mitomycin C—chemicals that induce DNA interstrand cross-links—and chromosomes are analyzed for evidence of chromatid breakage, which can produce characteristic abnormal chromosome formations (Figure 1C).

Figure 1 Examples of abnormal phenotypes in DNA-repair disorders. (A) Extensive skin cancer in xeroderma pigmentosum. (B) Accelerated aging in Werner syndrome—portraits of the same woman at age 13 (left) and age 56 (right). (C) Characteristic quadriradial and triradial chromosome formations in Fanconi anemia cells after treatment with mitomycin C. (A, courtesy of Himynameislax (CC BY-SA 3.0). B, from Hisama FM, Bohr VA, and Oshima J [2006] *Sci Aging Knowl Environ* 10:pe18. With permission from the AAAS (left) and the International Registry of Werner Syndrome (right). C, courtesy of Niall Howlett from Harney JA, Shimamura A and Howlett NG [2008] *Pediatr Health* 2:175−187. With permission from Future Medicine Ltd.)

(A)

(B)



(C)

4.2 POPULATION GENOMICS AND THE SCALE OF HUMAN GENETIC VARIATION

Genetic variation is caused by DNA sequence changes that can be sorted into different classes according to the underlying mechanisms and the scale. But all DNA changes can be classified into two broad categories:

1. ***Changes that do not affect the DNA content***. Here the number of nucleotides is unchanged. Quite often, for example, a single nucleotide is replaced by a different nucleotide. More rarely, multiple nucleotides at a time may be sent to a new location (by chromosome breakage and rejoining) without net loss or gain of DNA content; the great majority are balanced translocations and inversions resulting in chromosome breakage without net loss or gain of DNA.
2. ***Changes in copy number***. Here there is a net loss or gain of DNA sequence. At one extreme, abnormal chromosome segregation produces fewer or more chromosomes than normal, and therefore a change in the copy number of whole nuclear DNA molecules.

They are almost always harmful. At the other extreme are deletions or insertions of a single nucleotide. In between are copy number changes that range from altered numbers of sequences that may be short (specific oligonucleotide sequences, for example) or long (sequences extending over multiple Mb of DNA).

Overall, the most common DNA changes are those that change only a single nucleotide or a very small number of nucleotides. Small-scale changes like this (often called **point mutations**) may often have no obvious effect on the pheno-type; in that case they would be considered to be neutral mutations. That happens mostly because more than 90 % of our DNA has been poorly conserved during evolution and may have no or very little functional value to the cell. Small, and sometimes quite large, changes in this fraction of the genome seem to be without obvious effect.

DNA variants, polymorphisms, and human population genomics

Alternative forms of DNA produced by mutation are generally described as DNA **variants.** Until quite recently, it was usual to describe a common DNA variant, with a frequency > 0.01, as a **polymorphism**; DNA variants with frequencies < 0.01 were traditionally described as *rare variants*. (The 0.01 cut-off might seem arbitrary; it was initially proposed so as to exclude recurrent mutation.)

The general use of the term *polymorphism* has been declining, partly because of the arbitrary nature of the 0.01 cut-off, and partly because of ambiguity in how the term is used. In medical disciplines, for example, *polymorphism* is often used to denote any sequence variation that does not cause disease, whereas *mutation* is used to describe a disease-causing sequence variant.

In modern times, the term *polymorphism* is largely avoided in population genomics projects; instead, it is customary to use DNA *variants*; they are often classified as: common (frequency > 5 %); low frequency (from 0.5 %

to 5 %); and rare (< 0.5 %). And it is also now customary to describe a change to a single nucleotide as a *single nucleotide variant* (*SNV*).

The Human Genome Project delivered an artificial *reference* sequence for the human genome, a patchwork of partial genome sequences from multiple individual anonymized donors that were combined into a single sequence. To obtain detailed knowledge of human genetic variation, however, genome-wide sequences from multiple individuals (each with two nuclear genomes, a maternal and a paternal genome) need to be analyzed. Clearly, the greater the number of individual samples analyzed and the greater the fraction of the genome sequenced the more information is obtained. Analysis of very large numbers of genome sequences is important because rare genetic variants can be medically important.

Personal genome sequencing first became a reality in 2007–2008 (interested readers can find descriptions of the first two individual human genomes to be sequenced, one by laborious Sanger sequencing (PMID 17803354) and the other by rapid massively parallel DNA sequencing (PMID 18421352). The ability to sequence whole genomes rapidly ushered in the era of human *population genomics* (population-based genome sequencing). Detailed information on human genetic variation has rapidly become available and large-scale projects have been launched to correlate genotypes with phenotypes (see **Box 4.1**).

Structural variation as opposed to small-scale variation

The vast majority of DNA changes are errors of DNA replication and repair. They typically affect one or a very small number of nucleotides. Because of the predominance of small-scale mutations, the study of human genetic variation was very largely focused on this type of variation until quite recently.

Whole genome sequencing has shown that additional moderate to large-scale DNA changes (> 50 bp), which include the outcomes of specific types of DNA breakage and rejoining mechanisms, are also highly significant.

Such **structural variation** can involve very large changes, and although structural variants are rather infrequent, significantly more nucleotides across the genome are altered as a result of structural variation than as a result of small-scale mutations.

The borderline between small-scale genetic variation and structural variation is, of course, an arbitrary one. (In the past, structural variation used to be applied to sequences that were one kilobase or longer, but the modern tendency has been to include smaller variants as long as the sequence change involves more than 50 bp).

BOX 4.1 LARGE-SCALE HUMAN POPULATION GENOMICS AND GENOTYPE–PHENOTYPE CORRELATION PROJECTS

Once rapid personal genome sequencing became possible in 2008, the first human population genomics project was launched, the 1000 Genomes Project, the initial aim of which was to sequence 1000 individual genomes from 26 different human populations across the world. As well as getting more information generally on human genetic variants, a major goal was to compare the genetic diversity of different ethnic populations (substantial differences in genetic variants exist between different human populations and are important in explaining differential population-based susceptibility to many disorders).

Since then, there has been a plethora of human population genomics projects, some focused on sequencing whole **exomes** (concentrating on exons from protein-coding genes), and others on whole genome sequencing. The latter had the notable advantage of vastly increasing our knowledge of formerly neglected structural variation, as well as offering a wide range of single nucleotide variants outside the more intensively studied protein-coding regions.

Initially, different human population genomics projects often used diverse ways of analyzing the data. To improve efficiency, consortia were formed to aggregate the sequencing data from different projects, and re-

analyze the data in a common pipeline. In 2016, the Exome Aggregation Consortium (ExAC) catalogued genetic variation in the protein-coding parts of the genome from 60 000 people. And in 2020, the genome aggregation database consortium (gnomAD) reported on analysis of sequences from 125 748 human exomes and 15 708 whole genomes.

The large scale of the gnomAD study has been important for identifying rare variants, and the substantial number of whole genomes analyzed has provided important information on human noncoding DNA variation. For a short overview, see PMID 32461645; for seminal gnomAD research publications see under Further Reading.

GENOTYPE-PHENOTYPE CORRELATION PROJECTS

To maximize the value to medicine and health of the burgeoning data on human genetic variation, large projects have recently been developed with the aim of correlating genotypes with phenotypes. The UK Biobank Project has been a pioneering project in this respect, collecting deep genetic and phenotypic data on 500 000 individuals from across the UK. For the UK's 100 000 Genomes Project, genome sequences from 85 000 UK National Health Service (NHS) patients affected by a rare disease or cancer were sequenced by the end of 2018. That project is being expanded to 1 000 000 genomes, including those of the 500 000 UK Biobank volunteers, and a further expansion has been planned towards genome sequencing of 5 million UK individuals. The All of Us project organized by the US National Institutes of Health seeks to correlate geno-types in 1 000 000 volunteers with their health data. We consider genotype-phenotype correlations more fully in later chapters.

Small-scale variation: single nucleotide variants and small insertions and deletions

Base substitution is the most common type of point mutation. Two major classes of base substitution are recognized, as listed below:

1. a *transition* (a purine is replaced by another purine, or a pyrimidine by another pyrimidine)
2. a *transversion* (a purine is replaced by a pyrimidine, or a pyrimidine by a purine).

If, for example, an A were substituted, there are three possibilities: A ® C (transversion); A ® G (transition) or A ® T (transversion).

Base substitution can generate (**Figure 4.4**) a **single nucleotide variant (SNV)**. For example, at a defined position on a DNA molecule most sampled sequences might happen to have a G, but a minority might have a C. In many cases the minority variant occurs at a low frequency in the population and may be described as a rare variant, or even a *private variant* (as a result of very recent mutation). In some cases, however, a minority variant is present at a population frequency of 0.01 or more, a frequency that is too high to explain by recurrent mutation. In that case, the DNA variation has traditionally been described as a **single nucleotide polymorphism** or **SNP** [pronounced "snip"].)

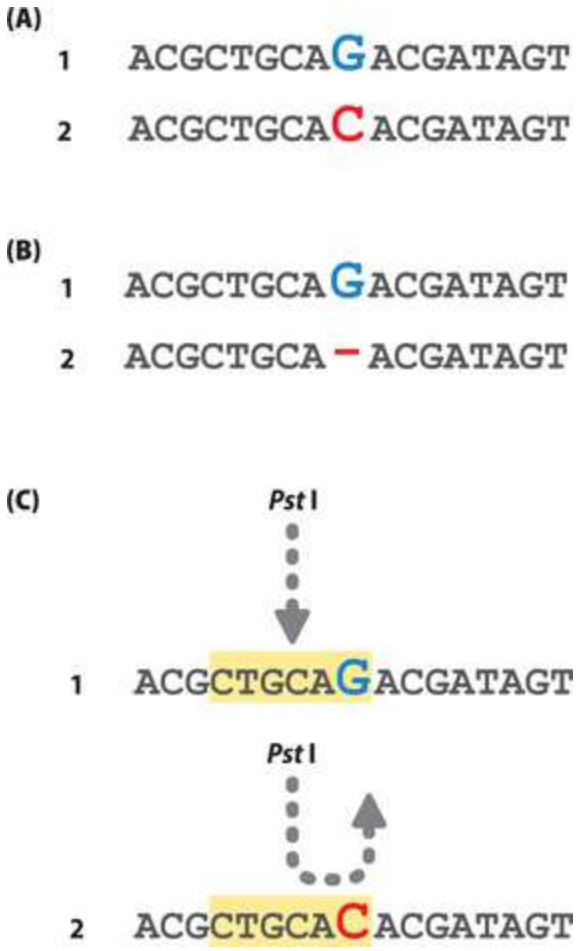


Figure 4.4 Classes of DNA variation affecting a single nucleotide position. (A) Single nucleotide variant (SNV) in which two variants differ by having a G or a C. (B) Insertion/deletion (*indel*) variation in which variant 1 has a G not present in variant 2. (C) Sometimes differences at a single nucleotide position can lead to the variable presence of a restriction site. Here variant 1 shown in (A) can be seen to have a recognition sequence (CTGCAG) for the restriction enzyme *PstI*; in variant 2 the equivalent sequence (CTGCAC) will not be cleaved by *PstI*. If the variants are common this would be an example of a *restriction fragment length polymorphism* (*RFLP*) that can conveniently be assayed by designing PCR primers to amplify the sequence containing it and then cutting the PCR product with *PstI*.

The pattern of single nucleotide variation in the human genome is nonrandom. Different DNA regions and different DNA sequences can undergo different mutation rates, and there is a large excess of C ® T substitutions in the human genome (see **Table 4.2**).

TABLE 4.2 SOME NONRANDOM FEATURES OF SINGLE NUCLEOTIDE VARIATION IN THE HUMAN GENOME

Feature	Description
Excess of transitions	The expected ratio of transitions to transversions is 1:2. In reality there is an excess of transitions because C → T transitions are unexpectedly frequent (see Figure 4.3)
Mutational bias	A general bias towards A-T base pairs (also observed in a wide range of other species)
SNP inheritance in germline DNA	Alternative nucleotides at SNP sites mark ancestral chromosomal segments that are common in the present-day population (see text)
Local suppression	Some regions of the genome, notably coding sequences, are subject to purifying selection to minimize harmful substitutions
Local enhancement	A higher rate in condensed chromatin (which is late replicating; possibly the condensed structure impedes access to the mismatch repair machinery)

Another reason for nonrandom variation comes from our evolutionary ancestry. Readers might reasonably wonder why only certain nucleotides should be polymorphic and be surrounded by stretches of nucleotides that only rarely show variants. In general, the nucleotides found at SNP sites are not particularly susceptible to mutation, and SNPs are stable over evolutionary time. Instead, the alternative nucleotides at SNP sites mark alternative ancestral chromosome segments that just happen to be common in the present-day population. As described in Chapter 8, using SNPs to define ancestral chromosome segments is important in mapping genetic determinants of disease. We cover methods for assaying specific single nucleotide changes in Chapter 11.

Small insertions and deletions

Some point mutations create DNA variants that differ by the presence or absence of a single nucleotide, or by a small number of nucleotides at a specific position. This is described as **in**sertion/**del**etion variation or *indel* for short—see Figure 4.4B. A subset of SNVs or indels leads to the gain or loss of a restriction site, in which case cutting the DNA with the relevant restriction nuclease can generate restriction fragment length polymorphism (RFLP; see Figure 4.4C).

Although indels could be considered to be copy number variants, the modern convention is to reserve the term **indel** to describe deletions or insertions of from one nucleotide up to an arbitrary 50 or so nucleotides (chosen because many massively parallel DNA sequencing methods produce quite short sequences and are often not suited to detecting deletions or insertion of greater than 50 nucleotides). The term **copy number variation** (**CNV**) is mostly used for changes in copy number of sequences that result in larger deletions and insertions, usually more than 100 nucleotides up to megabases.

The frequency of insertion/deletion polymorphism in the human genome is about one-tenth the frequency of single nucleotide substitutions. Short insertions and deletions are much more common than long ones. Thus, 90 % of all insertions and deletions are of sequences 1–10 nucleotides long, 9 % involve sequences from 11 to 100 nucleotides, and only 1 % involve sequences greater than 100 nucleotides. Nevertheless, because many of the last category involve huge numbers of nucleotides, CNV affects more nucleotides than single nucleotide substitutions.

Microsatellites and other variable number of tandem repeat (VNTR) polymorphisms

As described in Section 2.5, repetitive DNA accounts for a large fraction of the human genome. Tandem copies of quite short DNA repeats (1 bp to fewer than 200 bp) are common, and those with multiple tandem repeats are

especially prone to DNA variation. A continuous sequence of multiple tandem repeats is known as an array. Different organizations are evident and the repeated sequences are classified as belonging to three classes, according to the total length of the array and genomic location:

1. *microsatellite DNA* (array length: fewer than 100 bp long; widely distributed in euchromatin)
2. *minisatellite DNA* (array length: 100 bp to 20 kb; found primarily at telomeres and subtelomeric locations)
3. *satellite DNA* (array length is often from 20 kb to many hundreds of kilo-bases; located at centromeres and some other heterochromatic regions).

The instability of tandemly repeated DNA sequences results in DNA variants that differ in the numbers of tandem repeats, that is, variable number of tandem repeats (VNTR) polymorphism. Because microsatellite DNA arrays (usually called **microsatellites**) are frequently distributed within human euchromatin (roughly once every 30 kb) and are often highly polymorphic, they have been widely used in genetic mapping.

Because microsatellites have very short repeats (one to four base pairs long), microsatellite polymorphisms are sometimes known as *short tandem repeat polymorphisms (STRPs)*. They usually result in short insertions and deletions. But unlike SNPs (which almost always have just two alleles), microsatellite polymorphisms usually have multiple alleles (**Figure 4.5**).

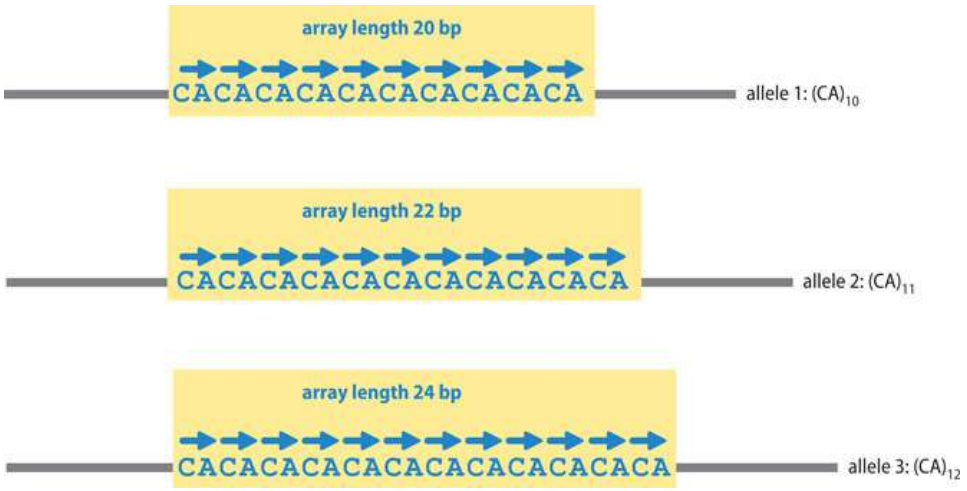


Figure 4.5 Length polymorphism in a microsatellite. Here, a microsatellite locus is imagined to have three common alleles that differ in length as a result of having variable numbers of tandem CA repeats. See **Figure 4.6** for the mechanism that gives rise to the variation in copy number.

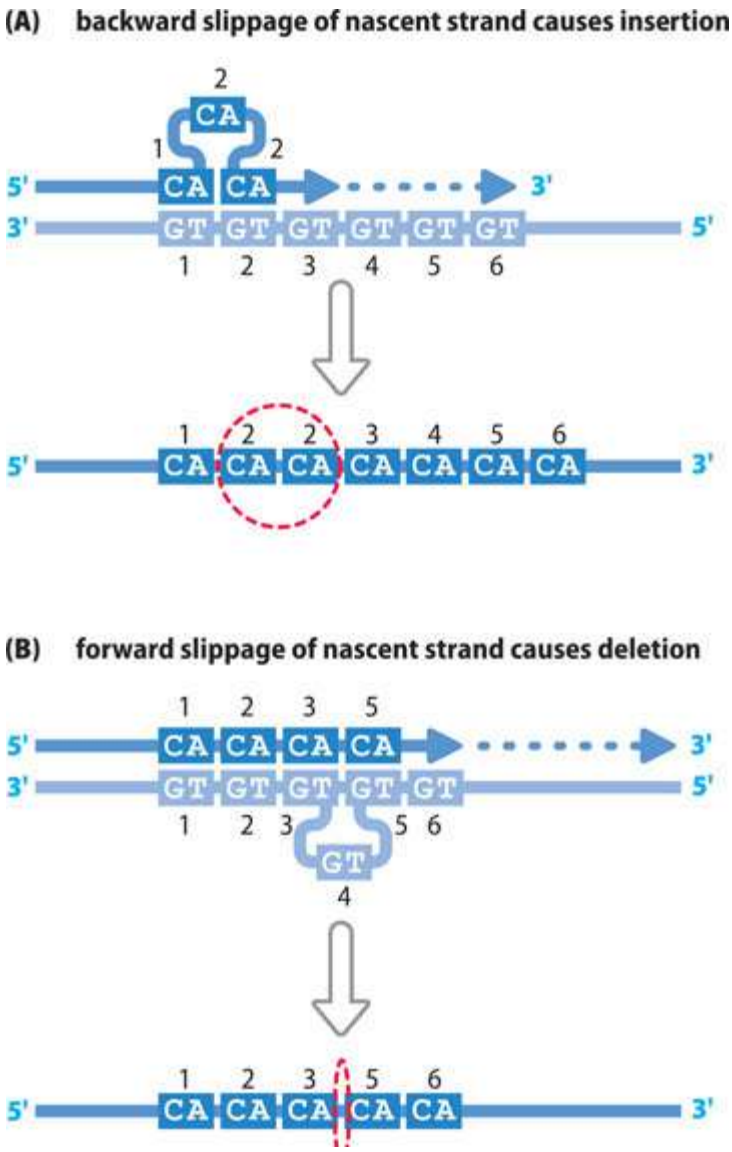


Figure 4.6 Microsatellite polymorphism results from strand slippage during DNA replication. The dark blue strand represents the synthesis of a new (nascent) DNA strand from the pale blue template DNA strand. During normal DNA replication, the nascent strand often partly dissociates from the template and then reassociates. When there is a tandemly repeated sequence, the nascent strand may mispair with the template strand when it reassociates, so that the newly synthesized strand has more repeat units (A) or fewer repeat units (B) than the template strand, as illustrated within the dashed red circles.

The variation in copy number arises as a result of **replication slippage**: during DNA replication, the nascent (newly synthesized) DNA strand slips

relative to the template strand so that the two strands are slightly out of alignment—see **Figure 4.6**.

Individual microsatellite polymorphisms can be assayed using PCR to amplify a short sequence containing the array, and then separating PCR products according to size by gel electrophoresis. Although extensively used in family studies and DNA profiling (to establish identity), it is not easy to automate assays of microsatellite polymorphisms, unlike for SNP assays.

Structural variation and low copy number variation

Until quite recently, the study of human genetic variation was largely focused on small-scale variation such as changes affecting single nucleotides and micro-satellite polymorphisms. We now know that variation due to moderately large-scale changes in DNA sequence is very common. Such **structural variation** can be of two types: balanced and unbalanced.

In balanced structural variation, the DNA variants have the same DNA content but differ in that some DNA sequences are located in different positions within the genome. They originate when chromosomes break and the fragments are incorrectly rejoined, but without loss or gain of DNA. That can involve inversions and translocations that do not involve change in DNA content (**Figure 4.7**).

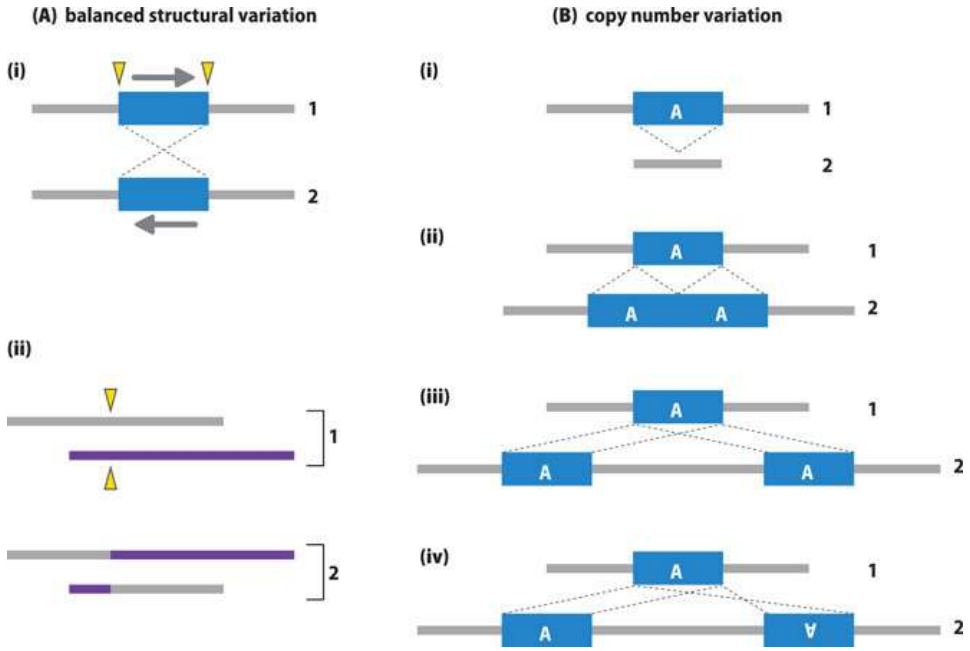


Figure 4.7 Structural variation and low copy number variation. The numbers 1 and 2 refer to alternative variants throughout. (A) Balanced structural variation involves large-scale changes that produce variants with the same number of nucleotides, including many inversions (i) and balanced translocations (ii). (B) Unbalanced structural variation includes unbalanced inversions and unbalanced translocations (not shown here) plus different types of low copy number variation (CNV). Copy number variants have different numbers of copies of a moderately long sequence (represented here by the box marked A). (i) CNV in which variants possess (1) or lack (2) a sequence, effectively a large-scale indel. This can result from an insertion (for example of a mobile element) or a deletion. (ii) CNV due to tandem duplication, effectively a large-scale VNTR that can sometimes have several copies, rather than just the one or two copies shown here. Sometimes additional insertion and inversion events can result in interspersed duplication with normal orientation of copies (iii) and interspersed duplication with inversion of a copy (iv).

In unbalanced structural variation, the DNA variants differ in DNA content. In rare cases in which a person has gained or lost certain chromosomal regions (as when a parent with a balanced reciprocal translocation passes one of the translocation chromosomes, but not the other, to a child), the gain or loss of substantial chromosomal segments

often results in disease. Unbalanced structural variation also includes commonly occurring CNV in which the variants differ in the number of copies of a moderately long to very long DNA sequence. Some CNVs such as this contribute to disease, but very many CNVs are commonly found in the normal population.

Copy number variation can take different forms. One form is effectively simple insertion/deletion variation on a large scale in which DNA variants either lack or possess a specific sequence (see Figure 4.7Bi). Other forms result from tandem duplication that may be complicated by subsequent insertion and inversion events (see Figure 4.7Bii–iv). In some CNVs, the DNA sequence that varies in copy number can include part of a gene sequence or regulatory sequence and sometimes multiple genes. As a result, some CNVs are important contributors to disease.

Taking stock of human genetic variation

The data from population-based genome sequencing projects indicate that single nucleotide changes are the most common type of genetic variation, accounting for close to 75 % of DNA changes. A study of 1092 individuals in 14 human populations by the 1000 Genomes Project Consortium reported a total of 38 million human SNPs (more than 1 per 100 nucleotides). However, the vast majority of these SNPs are rare in any population. Any one individual has just two haploid genomes and will be homozygous at many SNP loci. Personal genome sequencing shows that single nucleotide differences between the maternal and paternal genomes in one individual occur about once every 1000 nucleotides. Much of that variation falls outside coding sequences and mostly represents neutral mutations.

Structural variation is less common, accounting for close to one-quarter of mutational events and is dominated by CNV. Because CNV often involves very long stretches of DNA, however, the number of nucleotides involved in CNVs significantly exceeds those involved in SNVs. Various

databases have been established to curate basic data on genetic variation in humans (and other species)—see **Table 4.3**.

TABLE 4.3 GENERAL HUMAN GENETIC VARIATION DATABASES

Database	Description	Website
dbSNP	SNPs and other short genetic variations	http://www.ncbi.nlm.nih.gov/SNP/index.html
dbVar	genomic structural variation	http://www.ncbi.nlm.nih.gov/dbvar/
DGV		http://dgv.tcag.ca/
ALFRED	allele frequencies in human populations	http://alfred.med.yale.edu/alfred/index.asp

Databases focusing on mutations that relate to phenotypes and disease are described in Chapters 7, 8, 10 and 11.

4.3 FUNCTIONAL GENETIC VARIATION AND PROTEIN POLYMORPHISM

Until now we have focused on the different types of human genetic variation at the DNA level, and their origins. Most DNA variation is neutral, having no detectable effect on the phenotype, but a small fraction of DNA variation alters how gene products are made, causing disease. Surprisingly, however, a normal person carries an average of about 120 gene-inactivating variants, with about 20 genes being predicted to be inactivated in both alleles.

Such a level of gene inactivation might seem alarming. However, some predicted loss-of-function variants occur in exons that are variably used in transcripts. And quite a few of our genes do not carry out vital functions. For example, people with blood group O are homozygotes for an

inactivating mutation in the *ABO* gene. They fail to make an enzyme that transfers a monosaccharide group onto the H antigen (an oligosaccharide attached to certain lipids or proteins on the surface of some cell types, notably red blood cells). People with blood groups A, B, and AB do produce this enzyme. Effectively, therefore, people with blood group O make the H-antigen, and people with other blood groups make a modified H-antigen carrying an extra monosaccharide. Inactivating mutations are also common in some large gene families, such as the olfactory receptor super-family that we describe near the end of this section.

In this section we primarily consider how genetic variation is expressed at the level of gene products, notably proteins because they are overwhelmingly the major functional endpoint of our genes. (We have numerous RNA genes but either they assist protein synthesis directly or they are involved in gene regulation pathways ultimately affecting patterns of protein production. The great majority of molecular pathogenesis is ultimately due to abnormal protein expression).

Variation in protein sequences can occasionally result from recent gene duplication (as described below) but is usually due to variation at a single gene locus. In the latter case, the variation may result from changes at the level of DNA sequence or RNA sequence. Protein variants produced by changes in RNA sequence are usually described as **isoforms.** See **Table 4.4** for a summary.

TABLE 4.4 DIFFERENTWAYS OF PRODUCING VARIATION IN PROTEINS

Level	Mechanism	Examples
DNA (multilocus)	Gene duplication	Olfactory receptors (this section)
DNA (single locus)	Allelic variation	Numerous
	Post-zygotic DNA changes	Immunoglobulins,Tcell receptors (Section 4.4)

* We explain the underlying mechanisms in Chapter 6 when we consider gene regulation.

Level	Mechanism	Examples
RNA	Alternative splicing initiated from alternative promoters	p14and p16 from *CDKN2A* gene (Figure 6.8B)
	Alternative splicing causing variable possession of internal exons or sequence motifs	The +KTS and -KTS isoforms of the Wilms Tumor *WT1* gene
	RNA editing*	See Section 6.1
	Alternative polyA-addition*	Dystrophin Dp40 isoform

* We explain the underlying mechanisms in Chapter 6 when we consider gene regulation.

The vast majority of genetic variation has a neutral effect on the phenotype, but a small fraction is harmful

Functional DNA variants are primarily those affecting the function of genes, changing the structure of a gene product or altering the rate at which it is produced. Only a very small fraction of nucleotides in our DNA is important for gene function, however. Coding DNA sequences make up close to 1.3% of the human genome. Some additional DNA sequences make functional noncoding RNAs; others regulate gene expression in some way —at the transcriptional, post-transcriptional, or translational level.

Estimating how much of the genome is functionally important is not straightforward. The traditional way is to carry out cross-species comparisons to identify how much of the genome is subject to purifying selection to conserve functionally important sequences. Population-based human genome sequencing also offers insights into evolutionarily recent functional constraint. Current estimates suggest that perhaps a maximum of about 10 % of the genome is under functional constraint. Mutation at 90 % or more of our nucleotides may have essentially no effect.

Even within the small target of sequences that are important for gene function, many small DNA changes may still have no effect. For example,

many coding DNA mutations are silent: they do not change the protein sequence and would usually have no effect (unless they cause altered splicing—we show examples in Chapter 7). Single nucleotide changes in regulatory sequences or in sequences that specify functional noncoding RNA (ncRNA) may often also have no, or very little, effect. We know, however, little about the functional significance of changes in noncoding RNA; most studies have focused on coding DNA. Of course, harmful mutations also occur in functional DNA, and very occasionally mutations have a beneficial effect.

Different types of Darwinian natural selection operate in human lineages

A small fraction of genetic variation is harmful, and we consider the detail in other chapters, notably Chapters 7 and 10. Harmful mutations are subject to a type of negative selection called **purifying selection**: people who possess them will tend to have lower reproductive success rates and the mutant allele will gradually be eliminated from populations over several generations. Harmful DNA changes include many different types of small-scale mutations, both in coding DNA (resulting in changes in amino acid sequence) and noncoding DNA (causing altered splicing, altered gene regulation, or altered function).

In addition, structural variation can often have negative effects on gene function. Genes may be inactivated by balanced structural variation if breakpoints affect how they are expressed. Copy number variation can lead to a loss or gain of gene sequences that can be harmful because the levels of some of our gene products need to be tightly controlled, as explained in later chapters.

Occasionally, a DNA variant has a beneficial effect on the phenotype that can be transmitted to offspring. DNA variants like this become prevalent through **positive selection**. Here, individuals who possess the advantageous DNA variant may have increased survival and reproductive success rates;

the DNA variant then increases in frequency and spreads throughout a population.

Positive selection has occurred at different times in human lineages. It has been responsible for fostering different features that distinguish us from the great apes, notably human innovations in brain development and increased cognitive function. The great majority of the selected variants seem to occur in noncoding regulatory DNA and result in altered gene expression. As described below, positive selection is also important in response to microbial pathogens and to various alterations in our environment.

Positive selection in response to microbial pathogens

Human populations are subject to survival pressure from infectious diseases, notably ones that can develop through efficient transmission between humans to become pandemics. The Black Death plague in the mid-fourteenth century had mortality rates of 80–100 % and killed 25–50 % of European and Chinese populations; and 500 million people, one quarter of people on the planet at that time, were infected in the “Spanish” influenza pandemic of 1918 with a mortality rate of perhaps 10 %.

In the front line to protect us from diseases like this are HLA genes. They make proteins involved in recognizing viral and other foreign antigens in host cells and in activating T cells to recognize infected cells and counter the pathogen. The selection here is thought to favor heterozygosity at the key HLA loci. An individual then may often produce slightly different proteins at several HLA loci, increasing the chances of protective immune system responses. As a result, HLA proteins are the most polymorphic of all our proteins—we provide details in Section 4.4.

Adaptations to altered environments

Positive selection has also been responsible for various instances of adaptive evolution in human populations. After out-of-Africa migrations 50 000–100 000 years ago, modern humans settled in different geographic regions and were exposed to different environments, including living in latitudes with low levels of sunlight, or living at high altitudes. As shown in **Table 4.5** different adaptations developed in the migratory human populations.

TABLE 4.5 EXAMPLES OF GENETIC VARIANTS IN ADAPTIVE EVOLUTION IN HUMAN POPULATIONS

Altered environment	Adaptation and its effects	Associated genetic variants
Reduced sunlight (low UV exposure)	decreased pigmentation (decreased melanin in skin allows more efficient transmission of the depleted UV to a deep layer of the dermis—see text)	an *SLC24A5* variant (replacing the ancestral ALA at position 111 by THR) is prevalent in European populations (see Box 4.2)
High-altitude settlements (low O2 tension)	in Tibetan* populations lowered hemoglobin levels and a high density of blood capillaries provide protection against hypoxia	variants in *EPAS1,* a key gene in the hypoxia response
Malaria-infested environments	alterations in red blood cell physiology, affecting transmission of the mosquito-borne parasites *P. falciparum* or *P. vivax* and conferring increased resistance to malaria	pathogenic mutations** in *HBB* or *G6PD* for *P. falciparum* malaria; inactivating *DARC* variants that do not express the Duffy antigen* in *P vivax* malaria

Altered environment	Adaptation and its effects	Associated genetic variants
Lifelong intake of fresh milk	persistence of lactase production in adults, allowing efficient digestion of lactose	the -13910T allele about 14 kb upstream of the lactase gene, *LCT*
High levels of dietary starch	increased production of enzyme needed to digest starch efficiently	high *AMY1A* copy number (Figure 4.8)

Gene symbols are as follows: *SLC24A5,* solute carrier family 24, member 5; EPAS1, endothelial PAS domain protein 1; HBB, (β-globin gene; *G6PD,* glucose-6-phosphate dehydrogenase; *LCT,* lactase gene (converts lactose to galactose plus glucose); *AMY1A,* salivary α-amylase gene (converts starch into a mixture of constituent monosaccharides).

* Andean populations show different anti-hypoxia adaptations.

** Includes sickle-cell, thalassemia, and glucose-6-phosphate dehydrogenase deficiency mutations.

*** The Duffy antigen is a ubiquitously expressed cell surface protein required for infection of erthryocytes by *Plasmodium vivax.*

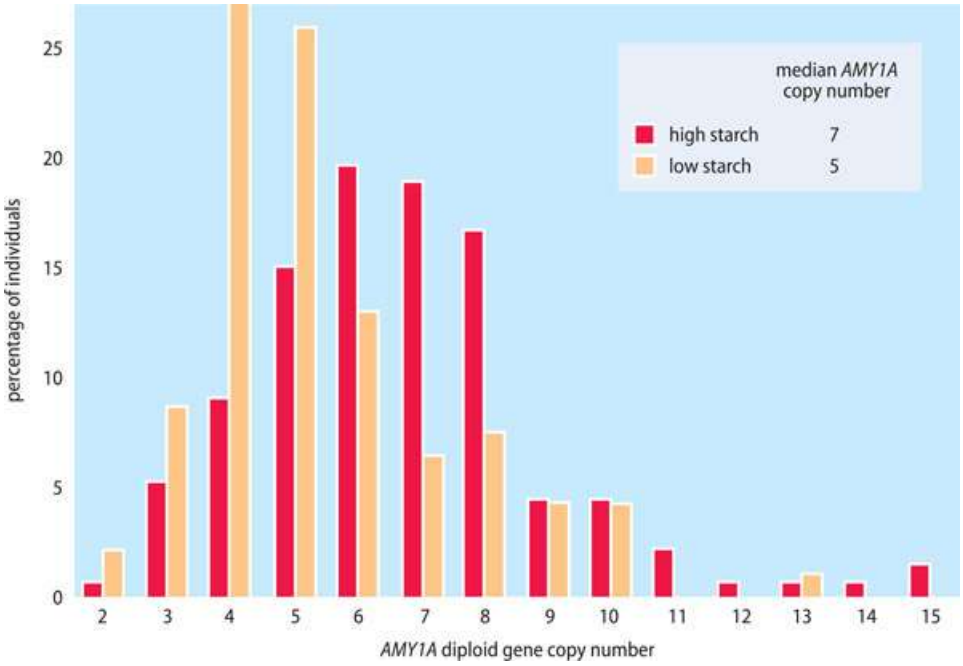


Figure 4.8 Recent acquisition of multiple genes encoding salivary α-amylase as an adaptation to high starch diets in some human populations. The graph illustrates

two points. First, the diploid copy number of the human salivary a-amylase gene *AMY1A* is quite variable (between 2 and 15 from this data set) but generally high (chimpanzees have a single copy of this gene). Secondly, individuals in populations that have high-starch diets have significantly more *AMY1A* gene copies than those in populations with low-starch diets. (From Perry GH et al. [2007] *Nat Genet* 39:1256–1260; PMID 17828263. With permission from Macmillan Publishers Ltd.)

Some adaptations also provide some protection against infectious diseases that are endemic in certain areas of the planet, notably mosquito-borne malaria. And as agriculture developed, DNA variants were selected in response to changes of diet (see Table 4.5).

Adaptations to local environments can involve downregulating a physiological function, as in reduced skin pigmentation in Europeans. Ultraviolet (UV) radiation in sunlight is needed for a photolytic reaction that occurs in a deep layer of the dermis, and this is the principal source of vitamin D3. Dark skins in equatorial populations protect skin cells from DNA damage caused by intense exposure to UV. Populations that migrated to northern latitudes were exposed to less UV, but the potentially reduced ability to make vitamin D3 was offset by an adaptation that reduced the amount of melanin, maximizing UV transmission through skin. The most significant contributor was a nonsynonymous change in the *SLC24A5* gene, resulting in the replacement of alanine at position 111 by threonine (A111T). The SLC24A5 protein is a type of calcium transporter that regulates melanin production, and the A111T change results in defective melanogenesis. The A111T variant became fixed in European populations as a result of what is called a *selective sweep* (**Box 4.2**).

BOX 4.2 RECENT STRONG POSITIVE SELECTION CAN LEAD TO A SELECTIVE SWEEP WITH LOCAL SUPPRESSION OF GENETIC VARIATION

Positive selection for an advantageous DNA variant can leave distinctive signatures of genetic variation in the DNA sequence. Imagine a large

population of individuals before positive selection occurs for some advantageous DNA variant on a region of, say, chromosome 22. If we were able to scan each chromosome 22 in the population before selection we might expect to find hundreds of thousands of different combinations of genetic variants (**Figure 1**).

Now imagine that an advantageous DNA variant arises by mutation on one chromosome 22 copy and then gets transmitted through successive generations. If the advantageous variant is subject to strong positive selection, people who carry it will have significantly higher survival and reproductive success rates. As descendants of the original chromosome 22 copy carrying the variant become more and more common, the selected DNA variant will increase in frequency to become a common allele (Figure 1B).

The entire chromosome 22 copy is not passed down as a unit: recombination will result in the replacement of some original segments by equivalent regions from other chromosome 22s. A short segment from the original chromosome 22 copy containing the favorable DNA variant and nearby "hitchhiking alleles" will increase in prevalence in a **selective sweep** (Figure 1B), but the segment will be slowly reduced in size by recombination.

A genomic region that has been subject to a recent selective sweep will demonstrate extremely low heterozygosity levels. The genomic region on chromosome 15 that contains the *SLC24A5* locus in Europeans provides a good practical example—see **Figure 2**.

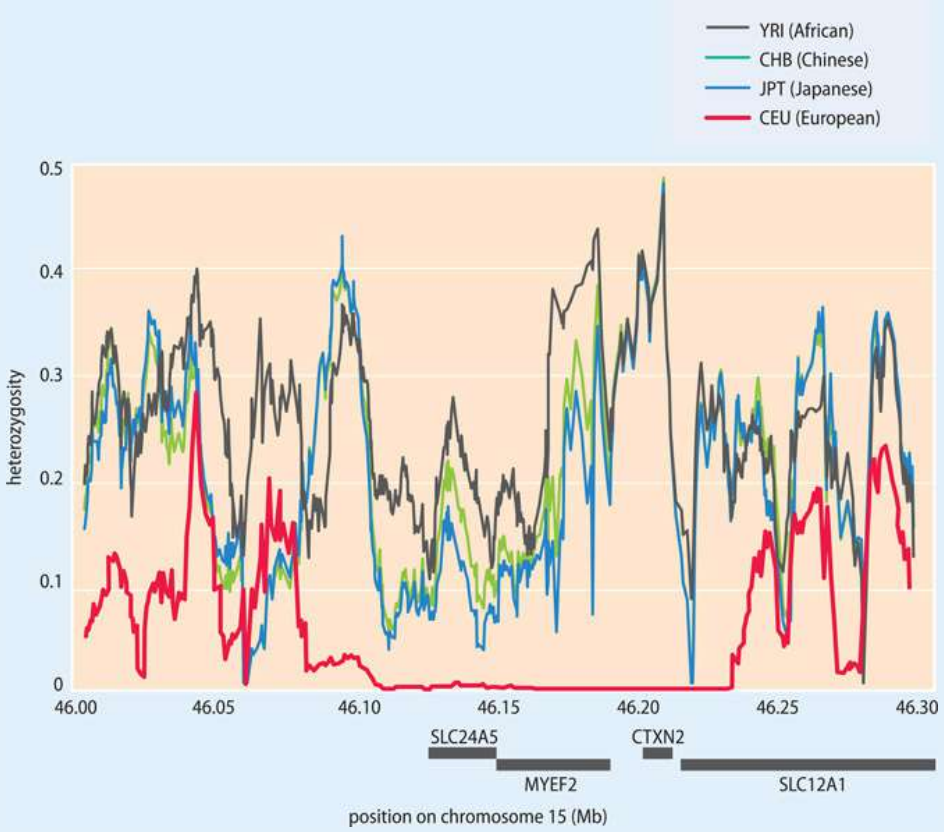


Figure 2 A strong selective sweep acting on an advantageous DNA variant in the *SLC24A5* gene in European populations. Heterozygosity levels in the region containing the *SLC24A5* gene on chromosome 15 were determined for a high density of common SNPs and averaged over 10 kb windows. The observed heterozygosity profiles for this chromosome region are unremarkable in African, Chinese, and Japanese populations. However, in the European population a strong selective sweep for a specific *SLC24A5* variant associated with reduced skin pigmentation has meant that almost all European chromosome 15s share a segment containing the favorable *SLC24A5* variant and hitchhiker alleles at the neighboring *MYEF2* and *CTXN2* loci. The result is a sharp decrease in heterozygosity for this chromosome region. (Adapted from Lamason RL, Mohideen MA, Mest JR et al. [2005] *Science* 310:1782–1786; PMID 16357253. With permission from the AAAS.)

Adaptations to living in malaria-infested regions have often involved increased frequencies of harmful alleles associated with certain blood

disorders, notably sickle-cell disease, thalassemia, and glucose-6-phosphate dehydrogenase deficiency in the case of *Plasmodium falciparum* malaria. Heterozygotes with mutant alleles associated with these diseases exhibit small changes in phenotype that make them comparatively resistant to malaria. **Balancing selection** is involved: the mutant heterozygotes have higher rates of reproductive success than both mutant and normal homozygotes (**heterozygote advantage**)—we consider the details in Chapter 5.

The development of agriculture brought significant changes to the human diet. The domestication of wheat and rice led to high-starch diets, and the domestication of cows and goats led to lifelong consumption of fresh milk in some human populations. Adaptive responses to high-starch diets and extended milk consumption both involved the increased production of enzymes required to metabolize starch or lactose (the major sugar in milk). But the adaptive genetic changes that permitted increased enzyme production were quite different as described below.

- *Gene amplification to permit enhanced starch metabolism.* Salivary a-amylase, the major enzyme needed to break down starch, is produced by the *AMY1A* gene. Our closest animal relative, the chimpanzee, has a single gene copy in the haploid genome but humans normally have multiple *AMY1A* genes. Individuals who take in a large amount of starch in their diets have a significantly higher *AMY1A* copy number and increased capacity to make salivary a-amylase than those used to low-starch diets (**Figure 4.8**).
- *Gene upregulation to permit enhanced lactose metabolism.* Like other mammals, most of the world's human population are lactose intolerant: the ability to digest lactose declines rapidly after weaning as levels of the enzyme lactase fall in the small intestine. In populations who had domesticated cows and goats, however, a cultural tradition developed of lifelong drinking of animal milk. Strong vertical transmission of this cultural practice led to selection for regulatory DNA variants allowing lifelong expression of the

lactase gene, *LCT* (lactase persistence). In each case mutations occur in a regulatory DNA region located about 14 kb upstream of the start codon.

Generating protein diversity by gene duplication: the example of olfactory receptor genes

Diverse forms of a protein – protein isoforms – can be generated by alternative mechanisms. Some occur at the level of post-transcriptional processing of an individual gene, including alternative splicing, alternative use of promoters, RNA editing and alternative polyA-addition (summarized in Table 4.4 above). Here we turn our attention to a different way of generating protein isoforms: gene duplication. Recall that gene duplication has been important in evolution in generating families of genes that can develop somewhat different functions. But in some cases gene duplication has been exploited to provide large numbers of protein isoforms that carry out the same basic function. The outstanding example is how gene duplication provides an extraordinary number of olfactory receptors.

The interaction of olfactory receptors on sensory neurons with odorants in the lining of the nose allows us to detect a huge number, possibly even one trillion, different smells, but there is pronounced variation between individuals in the ability to detect specific odors. As a single odorant may be recognized by multiple ORs, and one OR may recognize multiple odorants, there seems to be a combinatorial code of binding of different receptor so that ultimately, different odorants are represented as different combinations of activated ORs.

Perhaps not surprisingly, the olfactory receptor (OR) gene family is our largest protein-coding gene family (with ~400 OR genes plus ~450 OR pseudogenes) and this family demonstrates the greatest variation in gene content of any human gene family. In addition to the pseudogenes, alleles for deleterious mutations at functional OR gene loci are both common in the population and highly variable between individuals (**Figure 4.9**).

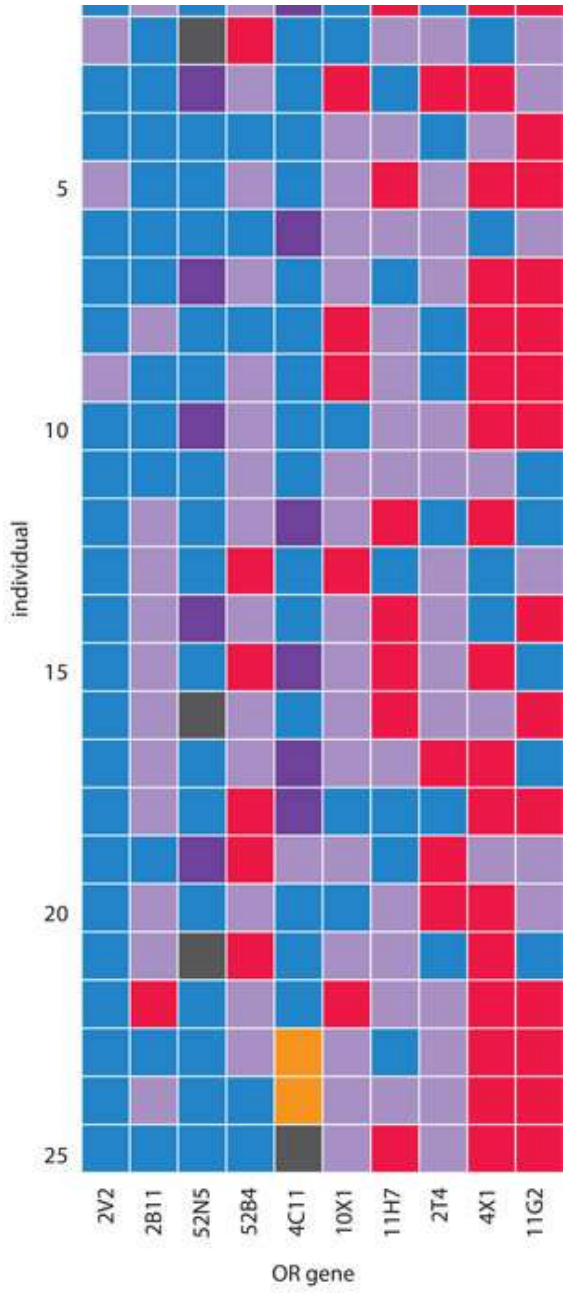

individual
5
10
15
20
25
2V2
2B11
52N5
52B4
4C11
10X1
11H7
2T4
4X1
11G2
OR gene
homozygous for deleterious SNP and deletion CNV
homozygous for deleterious SNP
homozygous for deletion CNV
heterozygous for deletion CNV
heterozygous for deleterious SNP
homozygous for intact allele

Figure 4.9 Common deleterious variants in human olfactory receptor (OR) genes. Colored squares represent recorded genotypes for different deleterious allele combinations at ten functional OR gene loci, for which both an intact and inactive allele are common in the population. Each column represents one of the ten OR genes studied; each row represents an individual person. (Data courtesy of Doron Lancet and Tsviya Olender, Department of Molecular Genetics, Weizmann Institute of Science, Rehovot.)

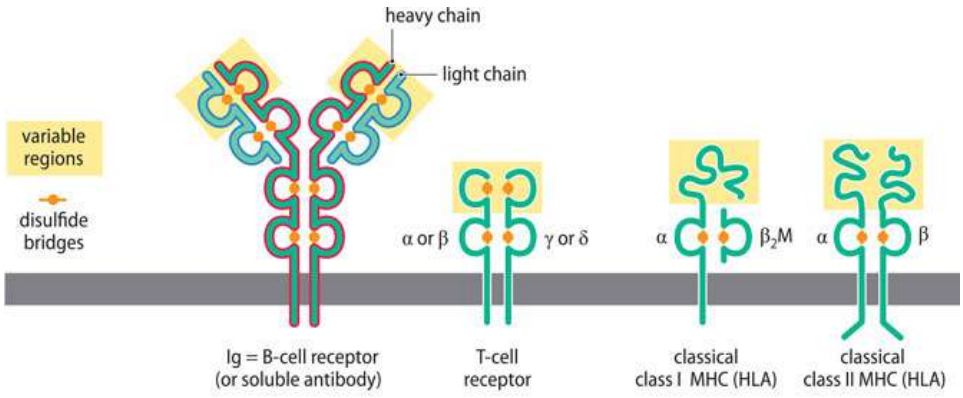
4.4 EXTRAORDINARY GENETIC VARIATION IN THE IMMUNE SYSTEM

Protein variants originating from a single gene locus through sequence changes at the DNA level are often rare. However, some proteins need to be able to recognize harmful foreign molecules introduced into the body that are subject to independent genetic control, and they can show very significant variation. The most extreme variation occurs in the case of immune system proteins working to recognize foreign molecules from microbial pathogens, ultimately leading to killing of microbes or virus-infected cells.

Pronounced genetic variation in four classes of immune system proteins

Our immune systems have a tough task. They are engaged in a relentless battle to protect us from potentially harmful microbial and viral pathogens. Not only must we be protected against a bewildering array of pathogens but, in addition, new forms of a pathogen can rapidly develop by mutation (in an effort to escape detection); that provides new challenges to which we must continuously adapt.

Four types of proteins are primarily involved, belonging to two broad classes as listed below and illustrated in **Figure 4.10**.



haploid number of human genes	3*	4*	3**	6**
number of different proteins	many millions in one person	millions in one person	up to 6 in one person; ~4000 in population	up to ~12 in one person; ~1200 in population
source of genetic variation	somatic recombination and hypermutation	somatic recombination	extensive polymorphism	extensive polymorphism

Figure 4.10 Extreme variation in four types of proteins needed to recognize foreign antigens. Immunoglobulins (Igs), T-cell receptors and MHC (major histocompatibility complex) proteins are heterodimers with similar structures: globular domains (maintained by intrachain disulfide bridges) and N-terminal *variable regions* that bind foreign antigens (but otherwise have a conserved sequence, known as *constant regions* in Igs). They are cell surface receptors (except that Igs in activated B cells become secreted antibodies). Only a few human genes encode an Ig or T-cell receptor, but nevertheless many different proteins are made because of special genetic mechanisms in B and T lymphocytes and because of selection for heterozygosity of HLA antigens. B_2M is β_2-microglobulin, the non-polymorphic light chain of class I HLA antigens. *It is estimated that we can make 10^7–10^8 different antibodies and close to the same number of different T-cell receptors. **Figure 1 of Box 4.3 on page 105 shows genes encoding the classical (= highly polymorphic) HLA proteins.

- *Immunoglobulins*. Expressed on the surface of B cells in the bone marrow or secreted as soluble immunoglobulins (antibodies) by activated B cells, their main task is to recognize and bind specific foreign antigens. They can bind and neutralize toxins released by

microbes, inhibit viruses from infecting host cells, and activate both complement-mediated lysis of bacteria and phagocytosis.

- *T-cell receptors*. Displayed on the surface of T cells, they work in cell-mediated immunity, along with proteins encoded by the major histocompatibility complex (MHC; known in humans as the HLA complex).
- *Classical class I MHC* (*HLA*) *proteins*. After synthesis they can recognize and bind foreign antigens within cells that have been infected by a virus or other cell surface pathogen. They transport the foreign antigens to the cell surface to be recognized by cytotoxic T lymphocytes (*antigen presentation*), after which the infected cells are killed by the cytotoxic T lymphocytes.
- *Classical class II MHC* (*HLA*) *proteins*. They are displayed on the surface of very few types of cells, notably immune system cells, and transport foreign intracellular antigens to the cell surface to be recognized by helper T lymphocytes.

As shown in Figure 4.10, there is an extraordinary variety of each of the above types of proteins, but their diversity arises by two different fundamental mechanisms. For MHC proteins, natural selection works to promote heterozygosity (if you have two protein variants at multiple HLA genes, you have a higher chance of recognizing and dealing with harmful antigens). But the extraordinary variation of immunoglobulins (Igs) and T-cell receptors (TCRs) comes from specialized mechanisms that are *programmed* to rearrange Ig genes in maturing B cells and TCR genes in maturing T cells.

A key point is that any *one* individual makes a huge variety of immunoglobulins and T-cell receptors because the rearrangement of Ig or TCR genes is *cell-specific*. Different B cells in a single individual can produce different immunoglobulins, and different T cells can produce different T-cell receptors. By contrast, the extensive variety of classical MHC proteins is apparent at a *population* level. Classical MHC proteins are highly polymorphic but a single person has a limited number of them,

having at most two alleles at each of a small number of polymorphic HLA loci.

Programmed and random post-zygotic genetic variation

As well as the genetic variation that we inherit from our parents, our DNA undergoes some changes as we develop from the single-celled zygote and throughout life. Post-zygotic genetic variation can involve mutations that occur randomly in all our cells. Therefore, although all our cells originate from the zygote, we are genetic **mosaics** who carry genetically different cells.

Much of the somatic genetic variation between the cells of an individual is due to copy number variants; mosaic patterns of copy number variations are a feature of human neurons, for example. Small-scale mutations also arise post-zygotically that often have no functional consequences. Whereas an inherited mutation will appear in all of our nucleated cells, a somatic mutation will only be present in the cell in which it arose plus any cell lineages that arise by cell division from the progenitor cell. Some somatic mutations give rise to disease if they occur at an early stage in development or result in abnormal tumor cell populations (we consider mosaicism for pathogenic mutations in **Box 5.3**).

In addition to random somatic mutations, programmed DNA changes are targeted to occur at immunoglobulin genes in the DNA of maturing B cells and at T-cell receptor genes in maturing T cells. We inherit from each parent just three immunoglobulin genes (*IGH, IGK,* and *IGL*) and four T-cell receptor genes (*TRA, TRB, TRD,* and *TRG*). However, the immunoglobulin genes in maturing B cells and the T-cell receptor genes in maturing T cells are programmed to undergo DNA changes *in a cell-specific way:* specific types of somatic DNA changes occur at these genes, but there is also a high degree of randomness so that the precise DNA changes vary from one B cell to the next B cell in the same individual, or from one T cell to the next. The net effect of these post-zygotic changes is to endow a single individual with huge numbers of

different immunoglobulin gene variants and of different T-cell receptor gene variants that can be pressed into service. Four mechanisms are responsible, as described in the next section.

Somatic mechanisms allow cell-specific production of immunoglobulins and T-cell receptors

Although a human zygote has a total of six immunoglobulin genes and eight T-cell receptor genes, somatic DNA changes in maturing B cells and T cells allow us to develop millions of different immunoglobulin (Ig) gene variants and millions of different T-cell receptor gene variants. Up to four mechanisms are involved, as described below.

Combinatorial diversity via somatic recombination

Each Ig and T-cell receptor gene is made up of a series of repeated gene segments that specify discrete segments of the protein, and different combinations of gene segments are used in protein production in different B cells or in different T cells of each individual. The different combinations of gene segments are made possible by somatic recombinations that occur in Ig genes in mature B cells and in T-cell receptor genes in mature T cells.

As an example, consider the gene segments that specify an Ig heavy chain. The variable region, which is involved in antigen recognition, is encoded by three types of repeated gene segment: V (encoding the first part of the variable region), D (diversity region), and J (joining region). The constant region defines the functional class of immunoglobulin (IgA, IgD, IgE, IgG, or IgM) and is encoded by repeated C gene segments (that have coding sequences split by introns). For each type of segment, the repeats are similar in sequence but nevertheless show some differences.

The first step in making an Ig heavy chain requires two sequential recombination events within the *IGH* gene of a maturing B cell. The end result is that one V gene segment, one D gene segment, and one J gene

segment are fused together to form a continuous VDJ coding sequence that will specify the variable region (Figure 4.11).

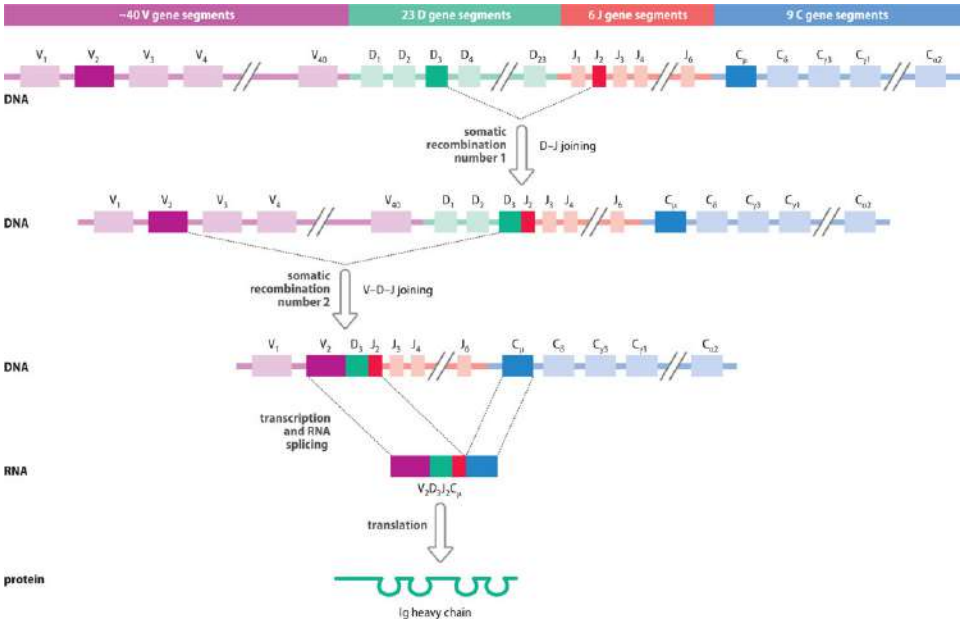


Figure 4.11 Somatic recombination in the *IGH* gene of B cells is used to make cell-specific immunoglobulin heavy chains. The human *IGH* gene has multiple but slightly different repeats for each of four types of gene segments: V (first part of variable region), D (diversity region), J (joining region), and C (constant region; although not shown here, each of the C gene segments has a coding sequence split by introns). An immunoglobulin heavy chain is made by bringing together coding sequences from one each of these four types of segments (shown here as filled boxes). Two sequential somatic recombinations produce first D–J joining, then a mature, functional VDJ coding sequence unit, which is effectively a large novel exon. In this example, the successful combination is V2D3J2, but the choice of combinations is *cell-specific*. Once a functional VDJ exon has been assembled, transcription is initiated starting with this exon and RNA splicing joins the VDJ coding sequence to coding sequences in the *closest* constant (C) region gene segment, in this case Cμ. Another type of somatic recombination (known as *class switching,* but not shown here) can change the position of C gene segments so that other C gene segments can be used instead of Cμ to give alternative classes of immunoglobulin.

Once assembled, a VDJ coding unit activates transcription. RNA splicing fuses the VDJ transcript to transcribed coding sequences within the nearest C gene segments, initially Cμ (see Figure 4.11) and then, through alternative splicing, either Cμ or Cq. The first immunoglobulins to be made by a B cell are membrane-bound IgM and then IgD. Subsequently, as B cells are stimulated by foreign antigen and helper T lymphocytes, they secrete IgM antibodies.

Later in the immune response, B cells undergo class-switching (also called isotype switching) to produce different antibody classes. Here, another type of somatic recombination positions an alternative C gene segment to be nearest to the J gene segments: either a C, Ce, or Ca gene segment to produce respectively an IgG, IgE, or IgA antibody.

The key point about the somatic recombination events that diversify the variable regions is that they occur randomly in each maturing B or T cell, respectively (with the proviso that one each of the different repeated gene segments are brought together). That is, the genetic variation is produced by *cell-specific* recombinations. The V2D3J2 combination in Figure 4.11 might occur in one maturing B cell, but a V4D2J9 unit or a V38D5J4 unit, for example, might be generated in neighboring B cells in the same person. The variable region of the T-cell receptor b chain is also formed by the same kind of VDJ recombination, but for both Ig light chains and T-cell receptor a chains a single VJ recombination is involved because their genes lack diversity gene segments.

Additional diversity generation

Two or three additional mechanisms are responsible for generating diversity in Igs and T-cell receptors as listed below. These mechanisms, together with V(D)J recombination, endow each of us with the potential of making many trillions of different antigen-binding sites, both for immunoglobulins and T-cell receptors. As required, individual B and T cells that enable succesful

recognition of foreign antigen are induced to proliferate to make identical clones with the same antigen specificity as the original cell.

- *Junctional diversity*. The somatic recombination mechanisms that bring together different gene segments in Ig or T-cell receptor genes variably add or subtract nucleotides at the junctions of the selected gene segments.
- *Protein chain combinatorial diversity*. Igs and T-cell receptors are heterodimers, and diversity is compounded by unique combinations of two unique protein chains. Note, however, that a B cell, for example, makes just one type of Ig. Although each diploid B cell has six Ig genes, in each B cell only one of the two *IGH* alleles is (randomly) selected to make a heavy chain (*allelic exclusion*) and only one of the four light chain genes in each B cell is ever used (a combination of *light chain exclusion*—to select either a k or l light chain—plus allelic exclusion).
- *Somatic hypermutation*. This mechanism applies only to Igs and is used to further increase variability in the variable region after somatic recombinations have produced functional VDJ or VJ units. When B cells are stimulated by a foreign antigen, an activation-induced cytidine deaminase is produced by the activated B cell that deaminates cytidine to uridine. The uridines are variably repaired by base excision repair (see above), and the end result is that multiple nucleotides in the variable region are mutated.

MHC (HLA) proteins: functions and polymorphism

The HLA complex is the human major histocompatibility complex (MHC). The latter name came from the observation that certain MHC genes are the primary determinants in transplant rejection. That, of course, is an artificial situation: the normal function of MHC genes is to assist certain immune system cells, notably helping T cells to identify host cells that harbor an intracellular pathogen such as a virus.

Some MHC genes—called classical MHC genes—are extremely polymorphic. They are subject to positive selection to maximize genetic variation (people who are heterozygous for multiple MHC loci will be better protected against microbial pathogens and have higher reproductive success rates). The classical MHC proteins are deployed on the cell surface as heterodimers (Figure 4.10). They serve to bind peptide fragments derived from the intracellular degradation of pathogen proteins and display them on the surface of host cells (**antigen presentation**) so that they can be recognized by T cells. Appropriate immune reactions are then initiated to destroy infected host cells. There are two major classes of classical MHC proteins, as detailed below.

Class I MHC proteins

Class I MHC proteins are expressed on almost all nucleated host cells. Their job is to help cytotoxic T lymphocytes (CTLs) to recognize and kill host cells that have been infected by a virus or other intracellular pathogen. When intracellular pathogens synthesize protein within host cells, a proportion of the protein molecules get degraded by proteasomes in the cytosol. The resulting peptide fragments are transported into the endoplasmic reticulum. Here, a newly formed class I MHC protein binds a peptide and is exported to the cell surface, where it is recognized by a CTL with a suitable receptor.

Because of cell-specific somatic recombinations (similar to those in Figure 4.11), individual CTLs make unique T-cell receptors that recognize *spe- cific* class I MHC–peptide combinations. If the bound peptide is derived from a pathogen, the CTL induces killing of the host cell. Note that a proportion of normal host cell proteins also undergo degradation in the cytosol and the resulting self-peptides are bound by class I MHC proteins and displayed on the cell surface. But there is normally no immune response (starting in early fetal life, CTLs that recognize MHC-self peptide are programmed to be deleted, to minimize autoimmune responses).

Class II MHC proteins

Class II MHC proteins are expressed in professional antigen-presenting cells: dendritic cells, macrophages, and B cells. These cells also express class I MHC proteins but, unlike most cells, they make co-stimulatory molecules needed to initiate lymphocyte immune responses.

Whereas class I MHC proteins bind peptides from *endogenous* proteins (those made within the cytosol, such as a viral protein made after infection of that cell), class II MHC proteins bind peptides derived from *exogenous* proteins that have been transported into the cell (by endocytosis of a microbe or its products) and delivered to an endosome, where limited proteolysis occurs. The resulting peptide fragments are bound by previously assembled class II MHC proteins and transported to the cell surface so that a helper T lymphocyte with an appropriate receptor recognizes a specific class II MHC–peptide combination. (Helper T cells have critical roles in coordinating immune responses by sending chemical signals to other immune system cells.)

MHC restriction

T cells recognize a foreign antigen only after it has been degraded and become associated with MHC molecules (***MHC restriction***). A proportion of all normal proteins in a cell are also degraded, and the resulting peptides are displayed on the cell surface, complexed to MHC molecules. MHC proteins cannot distinguish self from nonself, and even on the surface of a virus-infected cell the vast majority of the many thousands of MHC proteins on the cell surface bind peptides derived from host cell proteins rather than from virus proteins.

The rationale for MHC restriction is that it provides a simple and elegant solution to the problem of how to detect intracellular pathogens—it allows T cells to survey a peptide library derived from the entire set of proteins in a cell *but only after the peptides have been displayed on the cell surface*.

MHC polymorphism

MHC polymorphism is pathogen-driven: strong selection pressure favors the emergence of mutant pathogens that seek to evade MHC-mediated detection. The MHC has evolved two counterstrategies to maximize the chance of detecting a pathogen. First, gene duplication has provided multiple MHC genes that make different MHC proteins with different peptide-binding specificities. Secondly, many of the MHC genes are extraordinarily polymorphic, producing the most polymorphic of all our proteins (Table 4.6).

TABLE 4.6 STATISTICS FOR THE SIX MOST POLYMORPHIC HLA LOCI

HLA gene	***–A***	***–B***	***–C***	***–DPB1***	***–DQB1***	***–DRB***
Number of alleles or DNA variants	7354	8756	7307	1909	2193	3094
Number of protein variants	4302	5287	4042	1198	1386	2107

Data were derived from the European Bioinformatics Institute's IPD-IMGT/HLA database (release 3.47, January 2022). The statistics for these and additional loci are available at http://www.ebi.ac.uk/ipd/imgt/hla/about/statistics

The polymorphism of classical MHC proteins is focused on amino acids that form the antigen-binding pockets: different alleles exhibit different peptide-binding specificities. A form of long-standing balancing selection (also called overdominant selection) seems to promote MHC polymorphism. Heterozygosity is favored (presumably the ability to produce many different HLA proteins affords us greater protection against pathogens), and certain heterozygote geno-types seem to display greater fitness than others.

The balancing selection seems to have originated before the speciation event leading to evolutionary divergence from the great apes. HLA polymorphism is therefore exceptional in showing trans-species

polymorphism: a human HLA allele may be more closely related in sequence to a chimpanzee HLA allele than it is to another human HLA allele. For example, human HLA-DRB1*0701 and HLA-DRB1*0302 show 31 amino acid differences out of 270 amino acid positions, but human HLA-DRB1*0701 and its chimpanzee equivalent, Patr-DRB1*0702, show only 2 differences out of 270.

The medical importance of the HLA system

The HLA system is medically important for two principal reasons. First, the high degree of HLA polymorphism poses problems in organ and cell transplantation. Secondly, certain HLA alleles are risk factors for individual diseases, notably many autoimmune diseases and certain infectious diseases; other HLA alleles are protective factors, being negatively correlated with individual diseases.

Transplantation and histocompatibility testing

After organ and cell transplantation, the recipient's immune system will often mount an immune response against the transplanted donor cells (the graft), which carry different HLA antigens from those of the host cells. The immune reaction may be sufficient to cause rejection of the transplant (but corneal transplants produce minimal immune responses—the cornea is one of a few immune privileged sites that actively protect against immune responses in several ways, including having a much reduced expression of class I HLA antigens).

Bone marrow transplants and certain stem cell transplants can also result in graft-versus-host disease (GVHD) when the graft contains competent T cells that attack the recipient's cells. GVHD can even occur when donor and recipient are HLA-identical because of differences in minor (non-HLA) histocompatibility antigens.

Immunosuppressive drugs are used to suppress immune responses after transplantion, but transplant success depends largely on the degree of HLA

matching between the cells of the donor and the recipient. Histocompatibility testing (also called tissue typing) involves assaying HLA alleles in donor tissues so that the best match can be found for prospective recipients. The key HLA loci are the most polymorphic ones: *HLA-A, -B, -C, -DRB1, -DQB1,* and *-DPB1* (Table 4.6 and **Box 4.3**).

BOX 4.3 HLA GENES, ALLELES, AND HAPLOTYPES

HLA GENES

The HLA complex spans 3.6 Mb on the short arm of chromosome 6. The 253 genes in the complex include the 18 protein-coding HLA genes shown in **Figure 1**, ranging from *HLA-DPB1* (closest to the centromere) to *HLA-F*. Genes in the class I region make the heavy chain of class I HLA antigens (the non-polymorphic class I HLA light chain, b2-microglobulin, is encoded by a gene on chromosome 15); the class II region has genes encoding both chains of class II HLA antigens. The intervening region does not contain any HLA genes, but it does contain multiple genes with an immune system function and is sometimes referred to as the class III region.

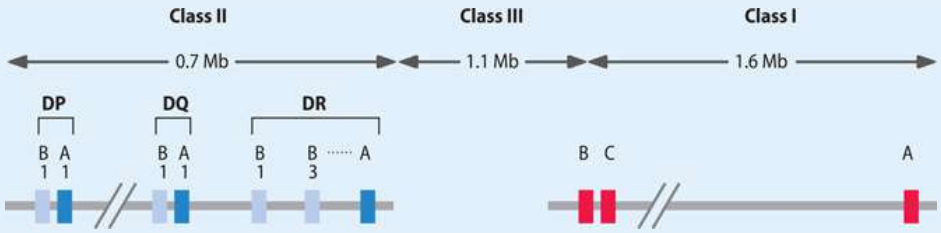


Figure 1 Classical (polymorphic) HLA genes within the HLA complex at 6p21.3. Genes in the class II HLA region encode a chains (dark shading) and b chains (pale shading) that pair up to form heterodimers within specific classes as indicated by horizontal bars above (DP, DQ, DR). Classical class I HLA genes encode a polymorphic class I a chain that forms a heterodimeric protein with the non-polymorphic b-microglobulin chain encoded by a gene on chromosome 15. Within the class I and class II HLA regions are several other non-polymorphic HLA genes and many HLA-related pseudogenes not shown here. The class III region includes certain

complement genes. Some additional genes with an immune system function are found within the HLA complex plus some functionally unrelated genes such as the steroid 21-hydroxylase gene.

HLA ALLELES

Because of their extraordinary polymorphism, alleles of the classical, highly polymorphic HLA genes have been typed for many decades at the protein level (using serological techniques with panels of suitably discriminating antisera). The number of alleles that can be distinguished in this way is very high, for example 28 HLA-A alleles, 50 HLA-B alleles, and 10 HLA-C alleles (called Cw for historical reasons; the "w" signifies workshop because nomenclature was updated at regular HLA workshops).

Serological HLA typing is still used when rapid typing is required, as in the case of solid organ transplants (in which the time between the chilling of an organ and the time it is warmed by having the blood supply restored needs to be minimized). However, much of modern HLA typing is performed at the DNA level, where very large numbers of alleles can be identified (see Table 4.6). The complexity means a rather cumbersome nomenclature for HLA alleles identified at the DNA level—see Table 1 for examples.

TABLE 1 HLAALLELE NOMENCLATURE

NOMENCLATURE	MEANING
HLA-DRB1	an HLA gene (encoding the β chain of the HLA-DR antigen)
*HLA-DRB1*13*	alleles that encode the serologically defined HLA-DR13 antigen
*HLA-DRB1*13:01*	one specific HLA allele that encodes the HLA-DR13 antigen

For more details see http://hla.alleles.org/

NOMENCLATURE	MEANING
*HLA-DRB1*13:01:02*	an allele that differs from *DRB1*13:01:01* by a synonymous mutation
*HLA-DRB1*13:01:01:02*	an allele that differs from *DRB1*13:01:01* by having a mutation outside the coding region
*HLA-A*24:09N*	a null allele related by sequence to alleles encoding the HLA-A24 antigen

For more details see http://hla.alleles.org/

HLA HAPLOTYPES

The genes in the HLA complex are highly clustered, being confined to an area that represents only about 2 % of chromosome 6. Genes that are close to each other on a chromosome are usually inherited together because there is only a small chance that they will be separated by a recombination event occurring in the short interval separating the genes. Such genes are said to be *tightly linked* (we consider genetic linkage in detail in Section 8.1).

A **haplotype** is a series of alleles at linked loci on an *individual* chromosome; haplotypes were first used widely in human genetics with reference to the HLA complex. See **Figure 2** for how haplotypes are established by tracking the inheritance of alleles in family studies. Note that because the HLA genes are very closely linked, recombination within the HLA complex is rare.

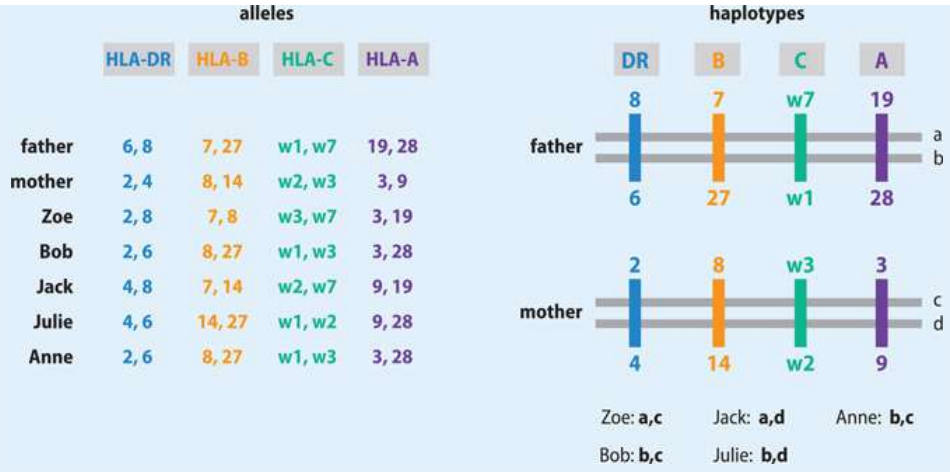


Figure 2 Deriving HLA haplotypes from family studies. Father, mother, and their three daughters, Zoe, Julie, and Anne, and two sons, Bob and Jack, have been tissue typed using serological reagents for four HLA antigens as shown at the left. By tracking which parental alleles have been passed on to individual children it is possible to deduce the parental HLA haplotypes. Father has one chromosome 6 with the HLA haplotype DR8, B7, Cw7, A19 (haplotype a) and another chromosome 6 with the HLA haplotype DR6, B27, Cw1, A28 (haplotype b). Similarly, mother has haplotypes c (DR2, B8, Cw3, A3) and d (DR4, B14, Cw2, A9). Father has transmitted haplotype a to Zoe and Jack, and haplotype b to Bob, Julie, and Anne. Mother has transmitted haplotype c to Zoe, Bob, and Anne, and haplotype d to Jack and Julie.

HLA disease associations

By displaying peptide fragments on host cell surfaces, HLA proteins direct T cells to recognize foreign antigens and initiate an immune response against cells containing viruses or other intracellular pathogens. Because HLA proteins differ in their ability to recognize specific foreign antigens, people with different HLA profiles might be expected to show different susceptibilities to some infectious diseases.

In autoimmune diseases, the normal ability to discriminate self-antigens from foreign antigens breaks down, and autoreactive T cells launch attacks

against certain types of host cells. Certain HLA antigens are very strongly associated with individual diseases, such as type 1 diabetes and rheumatoid arthritis; in general, genetic variants in the HLA complex are the most significant genetic risk factors that determine susceptibility to autoimmune diseases. Determining to what extent HLA variants are directly involved in the pathogenesis and how much is contributed by other variants that lie in the immediate vicinity of the HLA genes (and outside the HLA complex) is a major area of research—we consider HLA associations with individual diseases in some detail in Chapter 8.

SUMMARY

- The DNA in our cells accumulates changes over time (mutations) that usually have no significant effect on the phenotype.
- Some mutations adversely affect how genes work or are expressed; they can be associated with disease, and because at least some people carrying them have a lower reproductive fitness they tend to be removed from populations (purifying selection).
- Very occasionally, a mutation may result in some ben efit and may accumulate in frequency if it endows individuals with increased reproductive fitness (positive selection).
- Large-scale changes to DNA can result from abnormal ities in chromosome segregation and recombination. Smaller-scale changes typically result from unrepaired errors in DNA replication or unrepaired chemical attacks on DNA.
- DNA is damaged within living cells and organisms by various types of chemical attack that break covalent bonds in DNA or form inappropriate covalent bonds with bases. One or both DNA strands may be broken, bases or nucleotides may be

deleted, or inappropriate chemical groups may be covalently bonded to the DNA.

- Much of the chemical damage to DNA is caused by highly reactive chemicals produced naturally inside our cells.
- According to the type of chemical damage to DNA, different cellular pathways are used to repair a DNA lesion. Direct reversal of the damage-causing chemical steps is rare, and individual pathways often involve many molecular components.
- DNA variants often have low frequencies. More com mon variants, with a frequency of more than 0.01, are sometimes described as DNA polymorphisms.
- A single nucleotide variant (or polymorphism) involves the substitution of one nucleotide for another at a specific location. Nucleotide substitutions are nonrandom—for example, C ® T substitutions are particularly common in vertebrate DNA.
- An indel is a site where variants differ by lacking or possessing one or a few nucleotides.
- Some DNA variants differ by having different num bers of copies of a tandemly repeated DNA sequence, producing length variation. Microsatellite variants are DNA sequences that show small length differences as a result of having fewer or more tandem copies of a simple repeat sequence with between one and four nucleotides.
- Structural variation results from large-scale changes in DNA. In balanced structural variation, the variants do not differ in DNA content. In unbalanced structural variation, there is substantial length variation between variants that often occurs as a result of copy number variation for a long nucleotide sequence.

- In population-based genome sequencing, whole diploid genomes from multiple individuals are sequenced, providing comprehensive data on human genetic variation.
- Recent positive selection for genetic variants in dif ferent human populations has allowed adaptation to different local environments and to major dietary changes.
- Gene duplication is the basis of our diverse repertoire of olfactory receptors.
- To identify foreign antigens efficiently, each of us makes a huge variety of immunoglobulins and T-cell receptors. We inherit only three immunoglobulin and four T-cell receptor genes from each parent, but cell-specific somatic rearrangements in maturing B and T cells endow us with huge numbers of different immunoglobulin and T-cell receptor gene variants.
- Our most polymorphic proteins are produced by genes in the HLA complex (the human major histo-compatibility complex). HLA proteins recognize and bind peptides from processed foreign proteins and present them on cell surfaces so that they can be recognized by specific T-cell receptors.
- The extreme polymorphism of HLA proteins means that recipients of tissue or organ transplants often mount strong immune responses to the foreign tissue. Tissue typing seeks to find reasonable matches between HLA antigens expressed by donor tissue and prospective recipients.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

DNA damage and DNA repair

Barnes DE & Lindahl T (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet* 38:445–476; PMID 15568983.

Ciccia A & Elledge SJ (2010) The DNA damage response: making it to safe to play with knives. *Mol Cell* 40:179–204; PMID 20965415. [An authoritative review on both DNA damage and repair, including detailed tabulation of frequencies of different types of DNA damage and of inherited disorders of DNA damage responses/DNA repair.]

Rass U (2007) Defective DNA repair and neurodegenerative disease. *Cell* 130:991–1004; PMID 17889645.

Large-scale human population genomics and genotype-phenotype correlation projects

Bycroft C (2018) The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562:203–209; PMID 30305743.

Karczewski KJ (2020) The mutational constraint spectrum quantified from variation in 141 456 humans. *Nature* 581:434–443; PMID 32461654.

Nature's Genome Aggregation Database (gnomAD) Literature Collection available at https://www.nature.com/immersive/d42859-020-00002-x/index.html

The All of Us Research Program at www.allofus.nih.gov

The gnomAD genome aggregation database at https://gnomad.broadinstitute.org/

Structural variation, copy number variation, and indels

Alkan C (2011) Genome structural variation and genotyping. *Nature Rev Genet* 12:363–376; PMID 21358748.
Collins RL (2020) A structural variation reference for medical and population genetics. *Nature* 581:385–386; PMID 32461652.
Conrad DF (2010) Origins and functional impact of copy number variation in the human genome. *Nature* 464:704–712; PMID 19812545.
Mullaney JM (2010) Small insertions and deletions (INDELs) in human genomes. *Hum Mol Genet* 19:R131–R136; PMID 20858594.

Human mutation distribution and mutation rates

Campbell CD & Eichler EE (2013) Properties and rates of germline mutations in humans. *Trends Genet* 29:575–584; PMID 23684843.
Gonzalez-Perez A (2019) Local determinants of the mutational landscape of the human genome. *Cell* 177:101–114; PMID 30901533.
Lynch M (2010) Rate, molecular spectrum and consequences of human mutation. *Proc Natl Acad Sci USA* 107:961–968; PMID 20080596.

Functional variation, positive selection, and adaptive evolution

Fu W & Akey JM (2013) Selection and adaptation in the human genome. *Annu Rev Genomics Hum Genet* 14:467–489; PMID 23834317.
Mathieson I (2020) Human adaptation over the past 40 000 years. *Curr Opin Genet Dev* 62:97–104; PMID 32745952.
Olson MV (2012) Human genetic individuality. *Annu Rev Genomics Hum Genet* 13:1–27; PMID 22657391.
Prohaska H (2019) Human disease variation in the light of population genomics. *Cell* 177:115–131; PMID 30901534.

Post-zygotic genetic variation

Thorpe J (2020) Mosaicism in human health and disease. *Annu Rev Genet* 54:486–5120; PMID 32916079.

McConnell MJ (2013) Mosaic copy number variation in human neurons. *Science* 342:632–637; PMID 24179226.

Genetic variation in the immune system

Bronson PG (2013) A sequence-based approach demonstrates that balancing selection in classical human leukocyte antigen (HLA) loci is asymmetric. *Hum Molec Genet* 22:252–261; PMID 23065702.

Murphy K & Weaver C (2016) Janeway's Immunobiology, 9th ed. Garland *Science*.

Parham P (2021) *The Immune System*. 5th ed. W.W. Norton & Co.

Shiina T (2009) The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet* 54:15–39; PMID 19158813.

5

Single-gene disorders: inheritance patterns, phenotype variability, and allele frequencies

DOI: 10.1201/9781003044406-5

CONTENTS

Genes are functional units of DNA that make some product needed by cells, ultimately either the polypeptide chain of a protein or a functional noncoding RNA. In this chapter, however, we will view genes very largely

as abstract entities and we consider them within the context of single-gene disorders—diseases in which the genetic contribution is determined primarily by one gene locus. Although individually rare, single-gene disorders are important contributors to disease. Knowledge of single-gene disorders also provides a framework for understanding the more complex genetic susceptibility to common disease described in later chapters.

We look first at the patterns of inheritance of single-gene disorders and provide an introductory basis for estimating disease risk according to the inheritance pattern (we provide more advanced disease risk calculations within the context of genetic counseling in Chapter 11).

We also consider how genes affect our observable characteristics. The term phenotype may be used broadly to describe the observable characteristics of a person, an organ, or a cell. But geneticists also use the word phenotype in a narrower sense to describe only those specific manifestations that arise in response to the differential expression of just one or a small number of genes. These manifestations may be harmful, and we can talk of a disease phenotype.

When the observable manifestations are not disease-associated we normally refer to a character or **trait**, for example blue eyes or blood group O. We can measure and record aspects of the phenotype, such as anatomical and morphological features, behavior, or cognitive functions. Sophisticated laboratory procedures can be used to perform more extensive investigations at the physiological, cellular, and molecular levels.

Genetic variation—changes in the base sequence of our DNA—is the primary influence on the phenotype (identical twins are remarkably similar, after all). But it is not the only determinant of the phenotype: environmental factors also make a contribution. Gene expression can be regulated by various *epigenetic* mechanisms (which, unlike genetic mechanisms, are independent of the base sequence of DNA), and stochastic factors also make a contribution.

As we will see, there can be considerable complexity in the link between genetic variation and phenotype: even in single-gene disorders the phenotype is often variable in affected members of one family. Note that we

do not deal with the molecular basis of single-gene disorders here; that will be covered in later chapters, notably Chapter 7.

We end the chapter by generally looking at the factors that affect allele frequencies in populations, and then focusing on frequencies of disease alleles (which are important practically for calculating disease risk for some types of single-gene disorder). And we explain why some single-gene disorders are common but others are rare.

5.1 INTRODUCTION: TERMINOLOGY, ELECTRONIC RESOURCES, AND PEDIGREES

Background terminology and electronic resources with information on single-gene disorders

An individual gene or DNA sequence in our nuclear DNA has a unique chromosomal location that defines its position, its **locus** (plural **loci**). We can refer to the ABO blood group locus, for example, or the *D3S1563* locus (a polymorphic DNA marker sequence located on chromosome 3).

In human genetics an **allele** means an individual copy of a gene or other DNA sequence that is carried at a locus on a *single* chromosome. Because we are diploid, we normally have two alleles at any one chromosomal locus, one inherited from each parent: a maternal allele and a paternal allele). The term **genotype** describes the combination of alleles that a person possesses at a single locus (or at a number of loci). If both alleles are the same at an individual locus, a person is said to be **homozygous** at that locus and may be referred to as a homozygote. If the alleles are different, even by a single nucleotide, the person is said to be **heterozygous** at that locus, a heterozygote.

Although we are essentially diploid, men have two types of sex chromosomes, X and Y, which are very different in both structure and gene content. As a result, most DNA sequences on the X chromosome do not

have a direct equivalent (allele) on the Y chromosome, and vice versa. Men are therefore **hemizygous** for such loci (because they normally only have one allele). Women normally have two alleles at each locus on the X chromosome.

In humans any genetic character is likely to depend on the expression of a large number of genes and environmental factors. For some, however, a particular genotype at a *single* locus is the primary determinant, being both necessary and sufficient for the character to be expressed under normal circumstances. Such characters are often said to be **Mendelian**, but that implies that a chromosomal locus is involved; a more accurate term is monogenic (which takes into account both chromosomal loci and loci on mitochondrial DNA). Although collectively important, individual single-gene disorders are rare, and common genetic disorders depend on multiple genetic loci.

When a human monogenic disorder (or trait) is determined by a nuclear gene, the disorder (or trait) is said to be **dominant** if it is manifested in the heterozygote (who carries a normal allele and a mutant allele), or **recessive** if it is not. Sometimes two different phenotypes that result from mutations at a single gene locus can be simultaneously displayed by the heterozygote and are said to be co-dominant. For example, the AB blood group is the result of **co-dominant** expression of the A and B blood group phenotypes that are determined by different alleles at the *ABO* blood group locus. As described below, the inheritance of mitochondrial DNA is rather different, with important implications for associated phenotypes.

Various electronic resources provide extensive information on human single-gene disorders and characters (**Box 5.1**). GeneReviews® provides excellent summaries for many of the more common single-gene disorders that are accessible through the widely used PubMed system for electronic searching of biomedical research literature. We therefore often provide the eight-digit PubMed identifier (**PMID**) for relevant GeneReviews articles on single-gene disorders. The Online Mendelian Inheritance in Man (**OMIM**) database is comprehensive, and we provide six-digit OMIM database numbers for some disorders.

BOX 5.1 ELECTRONIC RESOURCES WITH INFORMATION ON HUMAN SINGLE-GENE DISORDERS AND UNDERLYING GENES

Some of the more comprehensive and stable resources are listed below. There are also many disease-specific databases; we describe some of these in Section 7.2.

GeneReviews (http://www.ncbi.nlm.nih.gov/books/NBK1116/; see PMID 20301295 for an alphabetic listing). This series of clinically and genetically orientated reviews of single-gene disorders is made available through NCBI's Bookshelf program. Individual reviews are assigned a PubMed identifier (PMID), an eight-digit number that in this case normally begins with 2030—for example, Huntington disease is at PMID 20301482. The series covers the most common single-gene disorders, and for listed disorders there is more clinical information than in OMIM (see below).

OMIM (http://www.ncbi.nlm.nih.gov/omim). The Online Mendelian Inheritance in Man database is the most comprehensive single source of information on human Mendelian phenotypes and the underlying genes. Entries have accumulated text over many years, and the early part of an entry may often reflect historical developments rather than current understanding.

Each OMIM entry has an identifying six-digit number in which the first digit indicates the mode of inheritance. The initial convention for the first digit was: 1, autosomal dominant; 2, autosomal recessive; 3, X-linked; 4, Y-linked; and 5, mitochondrial. However, the distinction between autosomal dominant and autosomal recessive was discontinued for new entries after May 1994. After that date all new entries for auto-somal traits and genes were assigned a six-digit number beginning with the number 6. See the review by McKusick (2007) under Further Reading for further details.

GeneCards® (http://www.genecards.org). A gene-centered database, this contains a large amount of automatically generated entries, mostly relating to specific human genes. It provides substantial biological information about each gene.

Investigating family history of disease and recording pedigrees

The extent to which a human disorder has a genetic basis can often be established by taking a family history. Medical records may be available to health service professionals for some family members; details of deceased family members and others who may be difficult to contact may be obtained by consulting more accessible family members.

A **pedigree** is a graphical representation of a family tree that uses the standard symbols depicted in **Figure 5.1**. Generations are often labeled with Roman numerals that increase from top to bottom of the page (toward the youngest generation). Individuals within each generation are given Arabic numerals that increase from left to right. An extended family covering many generations may be described as a kindred. A family member through whom the family is first ascertained (brought to the attention of health care professionals) is known as the *proband* (also called propositus—feminine proposita) and may be marked with an arrow.

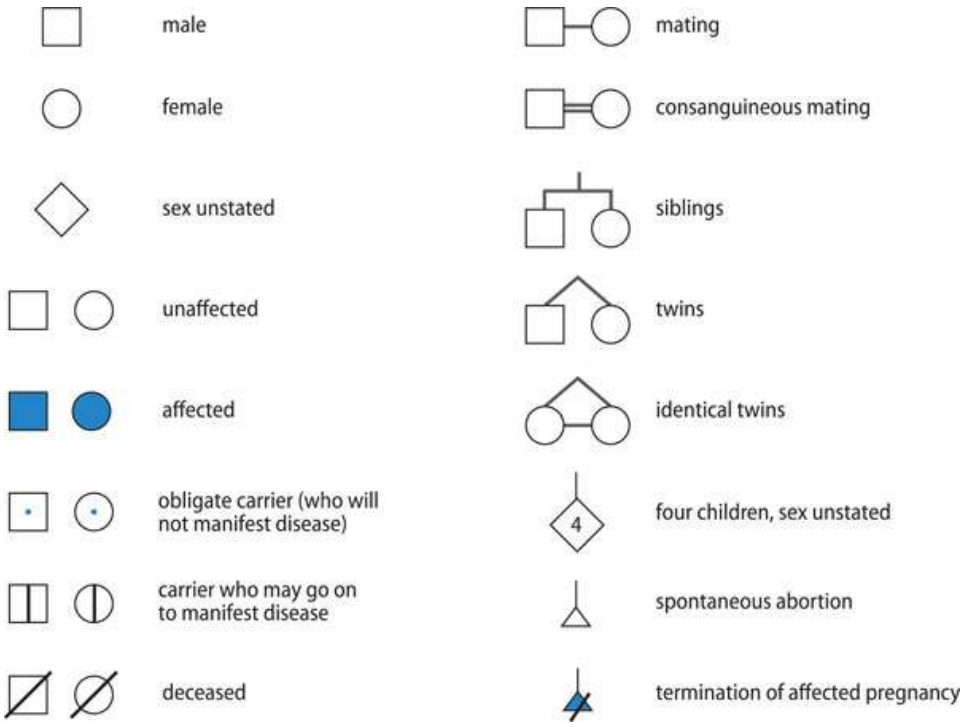


Figure 5.1 Pedigree symbols.

The term **sib** (sibling) is used to indicate a brother or sister, and a series of brothers and sisters is known as a sibship. According to the number of steps in the pedigree that links two family members, they may be classified relati*ves* of the first degree (parent and child; sibs); second degree (grandparent and grandchild; uncle/aunt and nephew/niece; half-sibs); third degree (first cousins), and so on. Couples who have one or more recent ancestors in common are said to be **consanguineous**.

5.2 THE BASICS OF MENDELIAN AND MITOCHONDRIAL DNA INHERITANCE PATTERNS

Mendelian characters are determined by chromosomal loci, either on an auto-some (human chromosomes 1 to 22) or on a sex chromosome (X or Y). Females are diploid for all loci (they have 23 pairs of homologous chromosomes). Males are different. Like females they have two copies of

each autosomal locus and of **pseudoautosomal** sequences found at the tips of the sex chromosomes (see below). However, they are hemizygous for the great majority of loci on the X and the Y (males have only one copy of the great majority of loci that are located on the X and the Y but outside the pseudoautosomal regions).

As a result of the above, there are five basic Mendelian inheritance patterns: autosomal dominant, autosomal recessive, X-linked dominant, X-linked recessive, and Y-linked (not Y-linked dominant or Y-linked recessive because males are never heterozygous for Y-linked sequences; the two Y chromosomes in rare XYY males are duplicates). In addition there is the unique pattern of inheritance of mitochondrial DNA mutations, which are substantial contributors to human genetic disease.

Autosomal dominant inheritance

A dominantly inherited disorder is one that is manifested in heterozygotes: affected persons usually carry one mutant allele and one normal allele at the disease locus. In autosomal dominant inheritance, the disease locus is present on an autosome (any chromosome other than the X or the Y), and so an affected person can be of either sex.

When an affected person has children with an unaffected person, each child would normally have a 50 % chance of developing the disease (the affected parent can transmit either the mutant allele or the normal allele). Affected persons often have an affected parent (see a typical example of autosomal dominant inheritance in **Figure 5.2**).

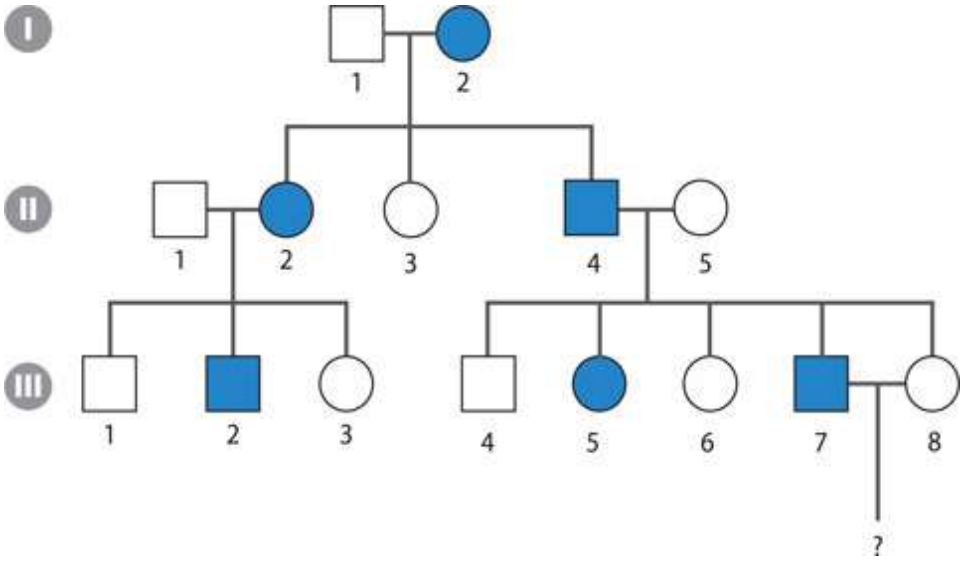


Figure 5.2 Pedigree showing autosomal dominant inheritance. Both sexes are affected and are equally likely to transmit the disorder. Affected individuals are typically heterozygotes (with one mutant allele and one normal allele) and usually have at least one affected parent. The question mark indicates the chance of having an affected child, which is 1 in 2 because one parent, III-7, is affected (there is a 50 % chance of transmitting the mutant allele and a 50 % chance of transmitting the normal allele).

Because the disorders are rare, affected individuals are almost always heterozygotes. Very occasionally, however, affected homozygotes are born to parents who are both affected heterozygotes. According to the effect of the mutation on the gene product, the affected homozygotes may show the same phenotype as the affected heterozygote. More commonly, affected homozygotes have a more severe phenotype than affected heterozygotes, as reported for conditions such as achondroplasia (PMID 20301331) and Waardenburg syndrome type I (PMID 20301703), or they have a much earlier age at onset of the disease, as in familial hypercholesterolemia (OMIM 143890).

In model organisms, a distinction is often made between different phenotypes seen in affected homozygotes and in affected heterozygotes (which are respectively called dominant and semidominant phenotypes in mice, for example). In human genetics, however, we refer to dominant phenotypes in

affected heterozygotes simply because affected homozygotes are so rarely encountered.

Autosomal recessive inheritance

A person affected by an autosomal recessive disorder can be of either sex and is usually born to unaffected parents who are heterozygotes (the parents would be described as asymptomatic **carriers** because they carry one mutant allele without being affected). Affected individuals carry two mutant alleles at the disease locus, one inherited from each parent. Assuming that both parents of an affected child are phenotypically normal carriers, the chance that each future child born to these parents is also affected is normally 25 % (the risk that one parent transmits the mutant allele is 1/2, so the risk that they both transmit the mutant allele to a child is 1/2 × 1/2 = 1/4).

Every one of us carries a single harmful allele at multiple loci associated with recessive phenotypes (carrying two such alleles can lead to disease, or even lethality in the prenatal period). When an autosomal recessive disorder is quite frequent, carriers will be common. In that case an affected child may often be born to two parents who carry different mutant alleles. The affected individual with two different mutant alleles would be described as a **compound heterozygote** (Figure 5.3A).

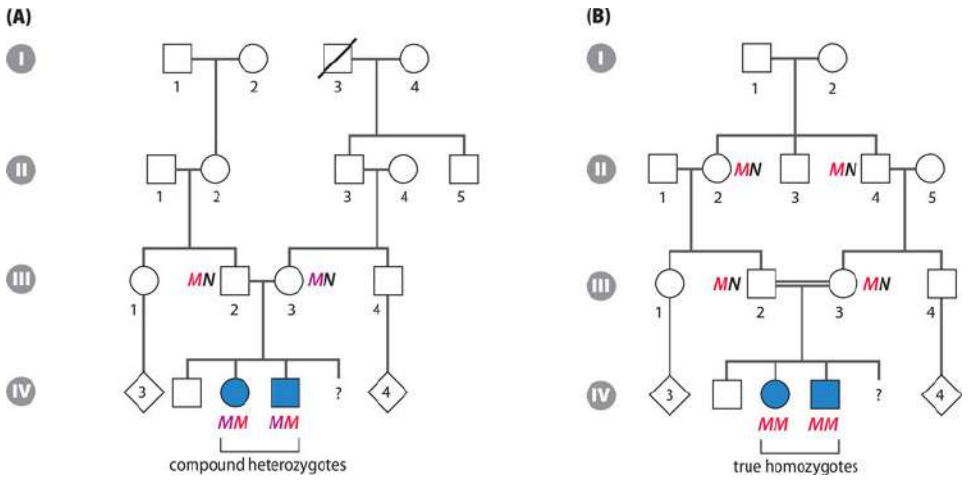


Figure 5.3 Pedigree showing autosomal recessive inheritance. (A) A pedigree for a common autosomal recessive disorder. The parents of the affected children in generation IV are carriers, with one normal allele (N) and one mutant allele (M). If they are not known to be related, they might well have different mutant alleles (shown by pink or red M) and the affected children would be compound heterozygotes. From the pedigree alone, we would not know who carried mutant alleles in generations I and II. For each subsequent child of III-2 and III-3, the risk of being affected is 1 in 4, irrespective of sex (each parent has a 50 % chance of transmitting the mutant allele, and the chance of inheriting both alleles is 1/2 × 1/2 = 1/4). (B) Involvement of consanguinity. Here we know that III-2 and III-3 are first cousins. They can be expected to be carriers, with one mutant allele (M) and one normal allele (N). We could infer that II-2 and II-4 inherited the same mutant allele (red M) from one parent (either I-1 or I-2). That means that III-2 and III-3 have the same mutant allele and their affected children will have inherited two identical mutant alleles and be true homozygotes. The chance that their fourth child will be affected (question mark) remains 1 in 4, irrespective of its sex.

Consanguinity

A feature of many recessive disorders, especially rare conditions, is that affected individuals often have two identical mutant alleles because the parents are close relatives; such couples are said to be *consanguineous*. In the example in Figure 5.3B the parents of the affected child in generation IV are first cousins, and they will have 1/8 of their genes in common by genetic descent (**Box 5.2** shows how these calculations are made). The two parents, III-2 and III-3, have each inherited the same mutant allele ultimately from the same common ancestor (in this case, a common grandparent, either I-1 or I-2).

BOX 5.2 CONSANGUINITY AND THE DEGREE TO WHICH CLOSE RELATIVES ARE GENETICALLY RELATED

Ultimately, all humans are related to one another, but we share the highest proportion of our genes with close family relatives. Mating between the most closely related family members (with 50 % of their genes in common, such as parent/child, and sibs) is very likely to result in homozygotes for recessive disease and is legally prohibited and/or socially discouraged in just about all societies. Cousin marriages can, however, be quite frequent in some communities from the Middle East, parts of the Indian subcontinent and other parts of Asia. Because cousins share a significant proportion of their genes, the offspring of cousin marriages can have a high degree of homozygosity with increased chance of being affected by a recessive disorder.

Because a child's risk of being homozygous for a rare recessive allele is proportional to how related the parents are, it is important to measure consanguinity. When one person is a direct descendant of another, the proportion of genes they have in common is $(1/2)^n$, where *n* is the number of generational steps separating the two. This gives: parent–child, 1/2 of genes in common; grandparent–grandchild, $(1/2)^2 = 1/4$ of genes in common; greatgrandparent–greatgrandchild, $(1/2)^3 = 1/8$ of genes in common.

CALCULATING THE COEFFICIENT OF RELATIONSHIP

The **coefficient of relationship** is the proportion of alleles shared by two persons as a result of common genetic descent from one or more recent (definable) common ancestors (or, more loosely, the proportion of genes in common as a result of common genetic descent). To calculate this, one considers paths of genetic descent linking the two individuals through *each* common ancestor in a family. A single generational step in such a path reduces the shared genetic component from the common ancestor by 1/2.

Consider the example in **Figure 1**. I-2 has had three children, a brother and sister who are sibs because they also have a common father, I-1, and their half brother, II-5. Half-sibs, such as II-3 and II-5, have a single

ancestor in common and so there is a single path connecting them to their common parent. So, the orange path connecting II-3 to II-5 via their common mother has two steps, making a contribution of 1/2 × 1/2 = 1/4 of genes in common.

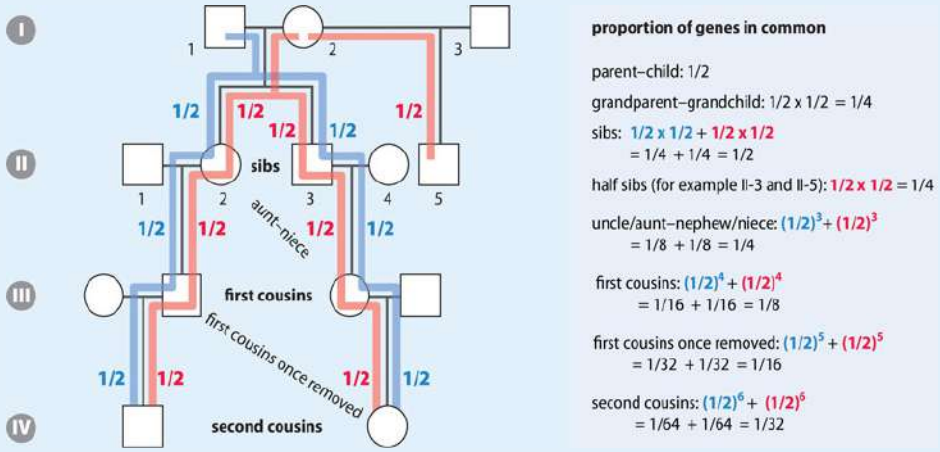


Figure 1 The proportion of genes in common between family members.

I-1 and I-2 are common ancestors for the sibs in generation II, for the first cousins in generation III and for the second cousins in generation IV. To calculate the coefficient of relationship for relatives linked by two or more common ancestors, we need to calculate the contributions made by each path and then sum them. Thus, for the first cousins in generation III, the green path that links them through their common grandfather, I-1, has four steps, making a contribution of $(1/2)^4 = 1/16$, and the orange path that links them through the common grandmother, I-2, also has four steps, making a contribution of $(1/2)^4 = 1/16$. Adding the two paths gives 2/16 or 1/8 of genes in common. More complicated inbreeding may mean that individuals have four or more recent common ancestors, but the principle is always the same: work out paths for each common ancestor and sum the contributions.

The **coefficient of inbreeding** is the probability that a homozygote has identical alleles at a locus as a result of common genetic descent from a recent ancestor. It is also the proportion of loci at which a person is expected to be homozygous because of parental consanguinity and is one-

half of the coefficient of relationship of the parents. So, if the parents are first cousins, the coefficient of inbreeding is 1/16. Note that even quite highly inbred pedigrees result in relatively moderate coefficients of inbreeding.

For rare disorders when there is doubt concerning the mode of inheritance, known parental consanguinity will strongly indicate autosomal recessive inheritance in a pedigree in which affected individuals have unaffected parents. But if consanguinity is not apparent, as in the pedigree in Figure 5.3A, alternative explanations are possible, as described below.

Disease-related phenotypes in carriers

Although carriers of an autosomal recessive disorder are considered asymptomatic, they may nevertheless express some disease-related trait that can distinguish them from the normal population. Take sickle-cell disease (OMIM 603903), for example. Affected individuals are homozygous for a b-globin mutation and produce an abnormal hemoglobin, HbS, that causes red blood cells to adopt a rigid, crescent (or sickle) shape. The sickle cells have a shorter life span that leads to anemia, and they can block small blood vessels, causing hypoxic tissue damage.

Carriers of the sickle-cell mutation are not quite asymptomatic. The sickle-cell allele produces a mutant b-globin that is co-dominantly expressed with the normal b-globin, and heterozygotes can have mild anemia (sickle-cell trait). However, under intense, stressful conditions such as exhaustion, hypoxia (at high altitudes), and/or severe infection, sickling may occur in heterozygotes and result in some of the complications associated with sickle-cell disease. Note that whereas sickle-cell disease is recessively inherited, the sickle-cell trait is expressed in the heterozygote and is therefore a dominant trait.

Sex-linked inheritance

In sex-linked inheritance, the inheritance patterns are controlled by genes that reside on the X and/or Y chromosomes. Before we go on to consider sex-linked inheritance, we need to take account of mechanisms that compensate for the variable number of sex chromosomes in humans (and other mammals): females have two X chromosomes but males have one X and one Y.

Having different numbers of chromosomes usually has severe, often lethal consequences—the loss of just one of our 46 chromosomes is lethal except for 45,X (Turner syndrome), and having an extra chromosome is usually lethal or results in a developmental syndrome such as trisomy 21 (Down syndrome). This happens because of problems with *gene dosage*: for some of our genes, the amount of gene product made must be tightly controlled (having one or three copies of these genes can be harmful because too little or too much product is made).

The sex difference regarding the Y chromosome is minimized by the conspicuous lack of genes on the Y. Most of the very few genes on the Y chromosome have male-specific functions, or they have an equivalent gene copy on the X (these X–Y gene pairs are mostly concentrated at the tips of the sex chromosomes in the pseudoautosomal regions).

X-chromosome inactivation

Unlike the Y chromosome, the human X chromosome has many hundreds of important genes. To compensate for having different numbers of X chromosomes in males and females, a special mechanism is needed: genes on one of the two X chromosomes in each female cell are silenced so that they do not produce any gene products (**X-inactivation**). Whereas males are *constitutionally* hemizygous for most genes on the X chromosome, X-inactivation means that at the *functional* level, females behave as if they were hemizygous for most genes on the X.

The X-inactivation mechanism is initiated after a cellular mechanism counts the number of X chromosomes in each cell of the early embryo. If

the number of X chromosomes is two (or more), all except one of the multiple X chromosomes is inactivated. Each such X chromosome is induced to form a highly condensed chromosome that is mostly transcriptionally inactive, known as a Barr body (**Figure 5.4**). Note that some genes, including genes in the pseudoautosomal regions, escape inactivation (we consider the mechanism of X-inactivation in Chapter 6).

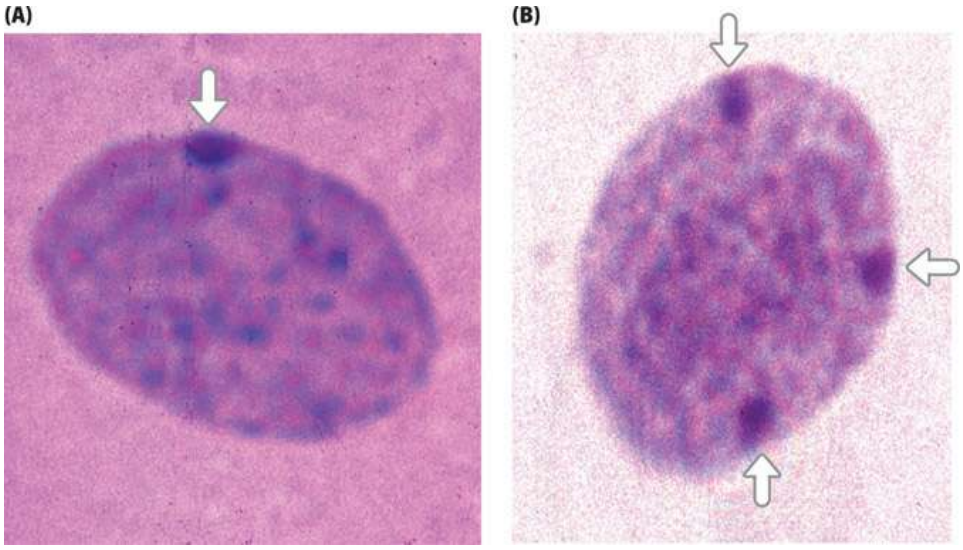


Figure 5.4 Barr bodies. (A) A cell from an XX female has a single inactivated X chromosome that forms a Barr body (arrow). (B) A cell from a 49,XXXXY male has one active X chromosome plus three inactivated X chromosomes that form Barr bodies (arrows). (Images courtesy of Malcolm Ferguson-Smith.)

In humans the initial decision to inactivate one of the two X chromosomes is randomly made in the preimplantation embryo, beginning at around the eight-cell stage; some cells inactivate the paternal X and others inactivate the maternal X. Once a cell has chosen which X to inactivate in the early embryo, however, that pattern of X-inactivation is continued in all descendant cells. Thus, a female who is heterozygous at a disease locus will be a genetic **mosaic**, containing cell clones in which the normal allele is expressed and clones in which the mutant allele is expressed. As described below, this has implications for the female phenotype in X-linked disorders.

X-linked recessive inheritance

In X-linked recessive disorders, affected individuals are mostly male, and affected males are usually born to unaffected parents. The mother of an affected male is quite often a carrier (and clearly so if she has affected male relatives). A distinguishing feature is that there is no male-to-male transmission because males pass a Y chromosome to sons (**Figure 5.5A**). However, a pedigree may *appear* to show male-to-male transmission when an affected man (with a condition such as hemophilia, for example) and a carrier woman produce an affected son (Figure 5.5B). The same parents could each potentially transmit a mutant X to produce an affected daughter.

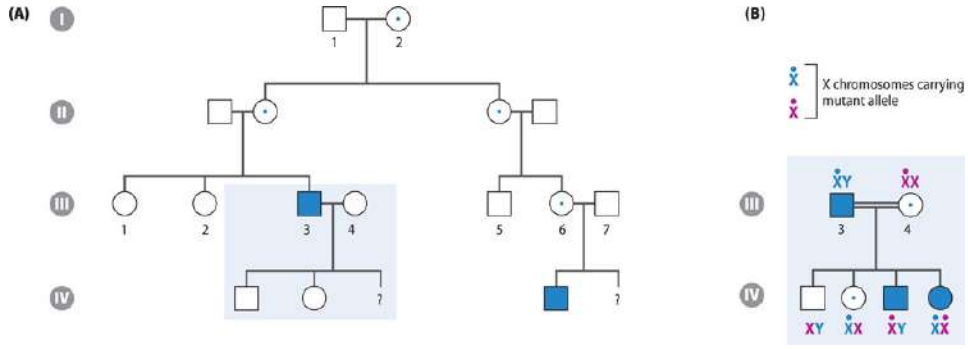


Figure 5.5 Pedigree showing X-linked recessive inheritance. (A) Affected males in generations III and IV have inherited (via female carriers) a common mutant allele from I-2. For each child of a carrier mother (such as III-6), the overall chance of being affected is 1 in 4, but this is sex-dependent: a son will have a 1 in 2 risk but a daughter will not be at risk (though she has a 1 in 2 risk of being a carrier). In the highlighted box, III-3 and III-4 have had two normal children, and the risk of having affected children would normally be very low (the father cannot transmit the mutant X allele to sons—he must transmit a Y—and any daughter will inherit a normal X from her mother). (B) The complication of inbreeding. Imagine that III-3 and III-4 were consanguineous and had the same mutant allele. We now have mating between an affected individual and a carrier, and there is a 1 in 2 chance that a child would be affected, irrespective of whether it was a boy or girl. The apparent male-to-male transmission is an illusion (the affected son has inherited a mutant allele from his mother, not from his father). The affected daughter is homozygous and although one

mutant allele will be silenced in each of her cells by X-inactivation, she does not have a normal allele.

In X-linked recessive disorders, female carriers with a single mutant allele can occasionally be quite severely affected and are known as *manifesting heterozygotes*. Because of X-inactivation, female carriers of an X-linked mutation are mosaics: some of their cells have the normal X chromosome inactivated and other cells have the mutant X inactivated, as seen most readily in skin disorders. Manifesting heterozygotes can occur by chance when most cells of a tissue critically important in disease development happen to have an inactivated X carrying the normal allele.

Manifesting heterozygotes can occasionally occur because of nonrandom X-inactivation. That can happen when there is some advantage in inactivating the normal X chromosome instead of the mutant X chromosome. For example, an X-linked disorder may manifest in a woman who has an X-autosome translocation in which the breakpoint on the X is the cause of the disorder. If the X-autosome translocation chromosome were to be inactivated, neighboring autosomal genes would also be silenced, causing gene dosage problems, and so the normal X is preferentially inactivated. Skewing of X-inactivation can often work in the other direction: some female carriers are asymptomatic because of nonrandom inactivation of the mutant X chromosome. We consider the mechanisms in Chapter 6.

X-linked dominant inheritance

As in autosomal dominant disorders, affected individuals with an X-linked dominant disorder can be of either sex and usually at least one parent is affected. However, there are significantly more affected females than affected males, and affected females typically have milder (but more variable) expression than affected males.

The excess of affected females arises because there is no male-to-male transmission of the disorder. All children born to an affected mother (and an unaffected father) have a 50 % chance of being affected, but an affected father with a single X chromosome will consistently have unaffected sons (they do not inherit his X chromosome), but his daughters will always be at risk because they will always inherit his affected X (Figure 5.6A gives an example pedigree).

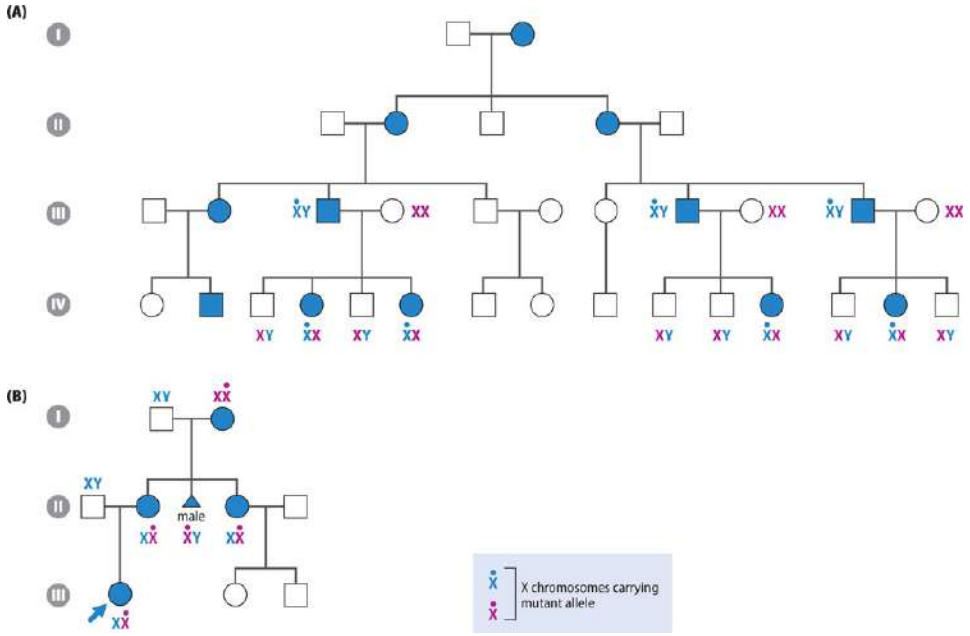


Figure 5.6 Pedigrees showing X-linked dominant inheritance. (A) Each child of an affected parent has a 1 in 2 chance of being affected. There is a 50 % chance that an affected female can transmit a mutant X allele to sons and to daughters; the risk is the same, irrespective of the sex of the child. However, the risk to the children of an affected father depends crucially on the sex of the child: every son would be expected to be unaffected; instead, the risk of being affected is focused on daughters (as shown for the three affected males in generation III who each have had children—the father must pass on a Y chromosome to each son and so does not transmit the mutant X, but must transmit the mutant X chromosome to each daughter). (B) X-linked dominant inheritance with early male lethality. This example shows four affected females in a three-generation family with incontinentia pigmenti that was followed up after the birth of the affected granddaughter (arrowed). An affected male in generation II had

spontaneously aborted. (Adapted from Minić S et al. [2010] *J Clin Pathol* 63:657–659; PMID 20591917. With permission from BMJ Publishing Group Ltd.)

The milder phenotype seen in affected females is a result of X-inactivation—the mutant allele is located on an inactivated X in a proportion of their cells. For certain X-linked dominant disorders, virtually all affected individuals are female: the phenotype is so severe in males that they die in the prenatal period, but the milder phenotype of affected females allows them to survive and reproduce. We illustrate this with the example of incontinentia pigmenti in Figure 5.6B; another disorder like this, Rett syndrome, is profiled in Section 6.3.

X-Y recombination and X-Y homology

In female meiosis, the two X chromosomes recombine like any pair of homologous chromosomes; in male meiosis, however, recombination between the X and Y chromosomes is very limited. The X and Y are very different in size, and pairing between the X and Y at meiosis is very limited.

Despite their considerable differences in size and gene content, the X and Y nevertheless have some short gene-containing regions in common, notably the pseudoautosomal regions located just before the telomere-associated repeats at the ends of both short and long chromosome arms (**Figure 5.7**). The pseudoautosomal regions are distinctive: they are the only regions of the X and Y that can pair up during male meiosis and undergo recombination like paired sequences do on homologous autosomal chromosomes (at each meiosis, there is an obligate X–Y crossover in the major pseudoautosomal region; recombination is less frequent in the minor pseudoautosomal region).

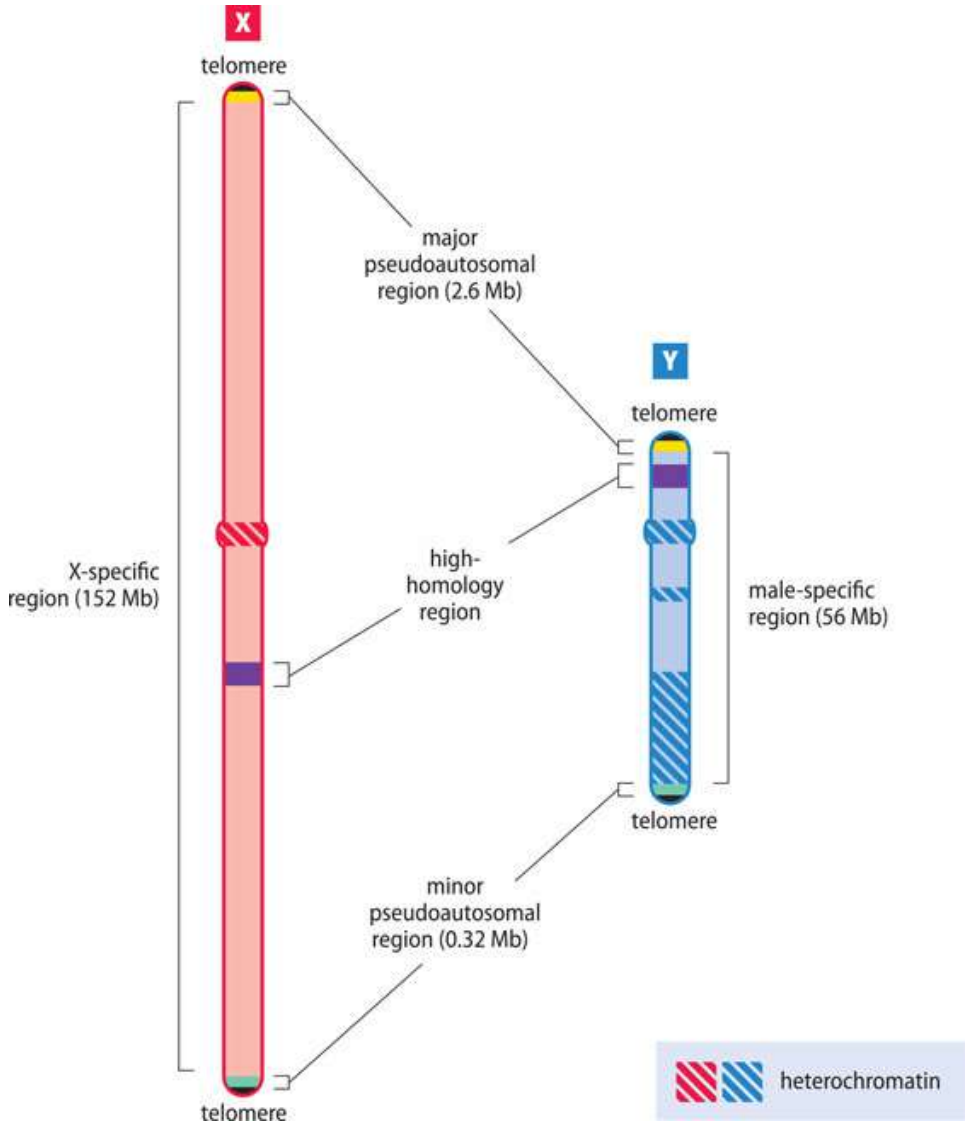


Figure 5.7 The human X and Y chromosomes: differences and major homology regions. The X and Y chromosomes differ greatly in size, heterochromatin content (much of the long arm of the Y is composed of heterochromatin), and DNA sequence. Colors indicate sequences that are X-specific (pink), Y-specific (blue), heterochromatic (hatched), or the major sequences shared by the X and Y chromosomes only (other colors). The major pseudoautosomal regions on the X and Y (yellow) are essentially identical, as are the minor pseudoautosomal regions on the long arms (green). The pseudoautosomal regions are involved in X–Y pairing and recombination in male meiosis. Note that the large central regions of the X and Y do not engage in recombination and are X-specific or Y-specific; the Y-specific region is also a *male-specific region* because it is not normally transmitted to females. As a result of an

evolutionarily recent X–Y transposition event there is also roughly 99 % sequence homology between certain sequences on Yp, lying close to the major pseudoautosomal region, and sequences at Xq21 (shown by purple boxes).

Outside the pseudoautosomal regions there is no recombination between the X and Y, and the remaining large central regions are X-specific and Y-specific regions. The X-specific region can engage in recombination in female meiosis, and sequences in this region can be transmitted to males or females; the Y-specific region is never involved in recombination and so is also called the *male-specific region*. The sequences in the X-specific and male-specific regions are very different, with just a few exceptions (see Figure 5.7 for an example).

Pseudoautosomal inheritance

As a result of recombination in male meiosis, the individual X–Y gene pairs in the pseudoautosomal regions are effectively alleles. An individual allele in these regions can move locations between the X and Y chromosomes and so is neither X-linked nor Y-linked; instead, the pattern of inheritance resembles autosomal inheritance (**Figure 5.8**).

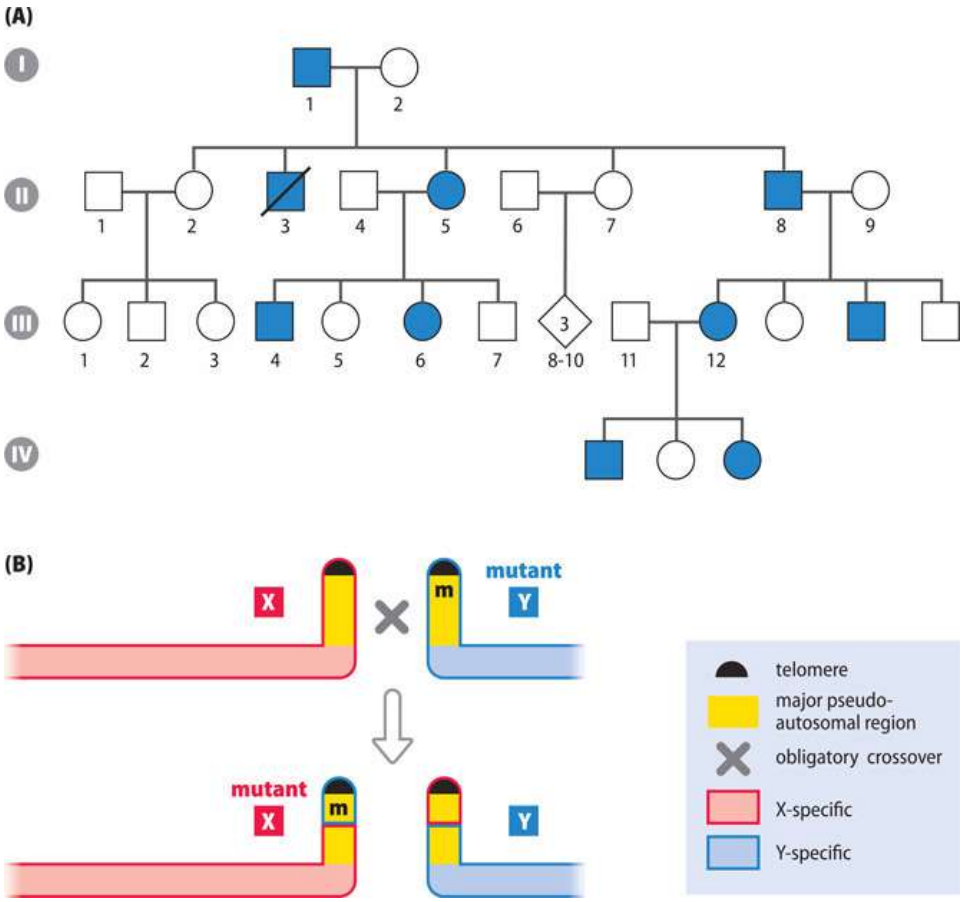


Figure 5.8 Pseudoautosomal inheritance and X–Y recombination. (A) In this pedigree, affected individuals are heterozygous for a mutation in the major pseudoautosomal region, and the disorder shows dominant inheritance. Affected females carry the mutation on an X chromosome that can be passed to both sons and daughters. Affected males pass a Y chromosome containing the mutant allele to affected sons but can also pass an X chromosome containing the mutant allele to affected daughters. This happens as a result of the obligatory X–Y recombination, which occurs within the major pseudoautosomal region. (B) When the crossover point occurs proximal to the mutant allele, the allele is transposed between the X and Y chromosomes.

There are few genes in the pseudoautosomal regions, so that few pseudoautosomal conditions have been described. However, mutations in *KAL* can cause Kallmann syndrome OMIM 308700, and the *SHOX* homeobox gene is a locus for two disorders. If one *SHOX* gene copy is damaged by mutation, the resulting heterozygotes have Leri-Weill

dyschondreostosis (OMIM 127300). Homozygotes with mutations in both *SHOX* genes have a more severe condition, Langer mesomelic dysplasia (OMIM 249700).

Y-linked inheritance

The Y-specific region of the Y chromosome is a nonrecombining male-specific region. Population genetics dictates that nonrecombining regions must gradually lose DNA sequences (as a way of deleting acquired harmful mutations because they cannot be removed by recombination). Over many millions of years of evolution, the Y chromosome has undergone a series of contractions and now has only 38 % of the DNA present in the X (the X and Y are thought to have originated as a homologous pair of autosomes that began to diverge in sequence after one of them acquired a sex-determining region). As a result of DNA losses, the male-specific region of the Y chromosome has few genes and makes a total of only 31 different proteins, most of which are involved in male-specific functions.

In Y-linked inheritance, males only should be affected and there should be exclusive male-to-male transmission. However, because of the lack of genes, Y-linked disorders are rare. Claims for some Y-linked traits, such as hairy ears (OMIM 425500), are now known to be dubious, but maleness is indisputably Y-linked. Interstitial deletions on the long arm of the Y chromosome are an important cause of male infertility (but infertile males are not normally able to transmit chromosomes unless conception is assisted by procedures such as intracytoplasmic sperm injection).

Matrilineal inheritance for mitochondrial DNA disorders

The mitochondrial genome is a small (16.5 kb) circular genome that has 37 genes (see Figure 2.11). It is much more prone to mutation than nuclear DNA, partly because of its proximity to reactive oxygen species (the mitochondrion is a major source of reactive oxygen species in the cell). As a result, mutations in mitochondrial DNA (mtDNA) are a significant cause

of human genetic disease. Tissues that have a high energy requirement—such as muscle and brain—are primarily affected in mtDNA disorders.

Individuals with a mitochondrial DNA disorder can be of either sex, but affected males do not transmit the condition to any of their children. The sperm does contribute mtDNA to the zygote, but the paternal mtDNA is destroyed in the very early embryo (after being tagged by ubiquitin), and a father's mtDNA sequence variants are not observed in his children. That is, inheritance occurs exclusively through the mother *(matrilineal* inheritance). An additional, common feature of mitochondrial DNA disorders is that the phenotype is highly variable within families (Figure 5.9).

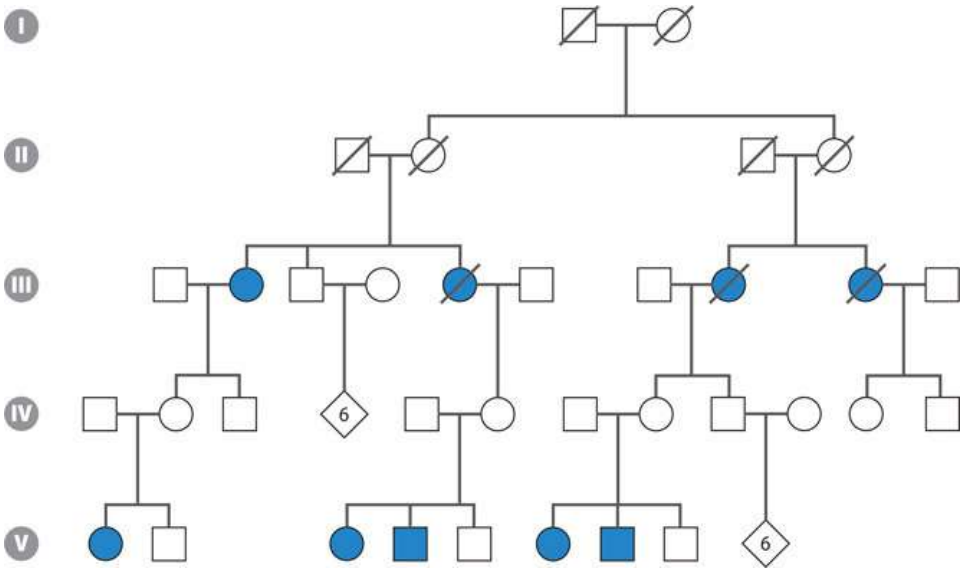


Figure 5.9 A pedigree illustrating matrilineal inheritance for a mitochondrial DNA disorder. Mitochondrial DNA disorders are transmitted by females only (because any mtDNA originating from the sperm is quickly degraded in the early embryo). However, an affected female can pass on the condition to both sons and daughters. A common feature of mtDNA disorders is incomplete penetrance, as shown here by the absence of clinical phenotypes in several individuals who must be gene carriers, including three clear carrier females in generation IV, each of whom were born to an affected mother and went on to produce affected children of their own (the females in generations I and II might also have been expected to be carriers of the mutant mtDNA). *One* cause of this intrafamilial variability is variable heteroplasmy. The mutation here was shown to be a nucleotide substitution in the mitochondrial 12S rRNA gene that was associated

with variable hearing loss. (From Prezant TR et al. [1993] *Nat Genet* 4:289–294; PMID 7689389. With permission from Macmillan Publishers Ltd.)

Variable heteroplasmy and clinical variability

Each cell contains multiple mitochondria, and there are often several hundred to thousands of mtDNA copies per cell. In some affected persons, every mtDNA molecule carries the causative mutation (homoplasmy), but affected individuals frequently have cells with a mixed population of normal and mutant mtDNAs (**heteroplasmy**). The clinical features depend mostly on the proportion of mutant to normal mtDNA molecules in the cells of tissues with high energy requirements.

Although a human egg cell is haploid for nuclear DNA, it contains more than 100 000 mtDNA molecules. A heteroplasmic mother can give rise to children who differ widely from her and from each other in the ratio of mutant to normal mtDNA molecules in their tissues (variable heteroplasmy). As a result, there can be very significant clinical variability between affected members of the same family.

To explain rapid shifts in heteroplasmy that occur over only one generation, the mitochondrial genetic bottleneck hypothesis envisages that, during early development, germline cells pass through a bottleneck stage in which they contain very few mtDNA molecules. By chance, germline cells at this stage may have a much higher or much lower proportion of mutant mtDNA molecules than the somatic cells. As a result, a heteroplasmic mother could give rise ultimately to eggs with a much higher or much lower proportion of mutant mtDNA molecules than are present in her affected tissues. We consider this in more detail in Section 7.6.

Another contributor to variable heteroplasmy is the rapid evolution of a mtDNA variant within an individual. Mutant mtDNAs that have a large deletion or a large duplication can evolve rapidly, so that different tissues or even the same tissue at different times may show different distributions of the mtDNA variant.

5.3 UNCERTAINTY, HETEROGENEITY, AND VARIABLE EXPRESSION OF MENDELIAN PHENOTYPES

Section 5.2 dealt with the different modes of inheritance for phenotypes that are determined principally by single genes. Some complications were covered, including the effects of X-inactivation in females and hemizygosity in males, occasional differences between homozygotes and heterozygotes for autosomal dominant disorders, the occasional expression of disease symptoms in carriers of an autosomal recessive disorder, the mimicking of autosomal inheritance by genes in the pseudoautosomal regions of the X and Y chromosomes, and the unique features of mitochondrial DNA inheritance.

In this section we discuss broader complications that relate to uncertainty of mode of inheritance, and difficulties posed by heterogeneity in the links between DNA variation and phenotypes. In addition, we consider how affected individuals within a single family can show variable phenotypes for Mendelian disorders.

Difficulties in defining the mode of inheritance in small pedigrees

Many families are small and may have only a single affected person. If the disorder is rare and we do not know the underlying disease gene, how can we work out the mode of inheritance? Knowing the mode of inheritance is important in genetic counseling (calculating the risk of having a subsequent affected child is made on that basis). Unless a disease gene has been identified and screened for mutations, however, the mode of inheritance inferred from examining the pedigree should be regarded simply as a working hypothesis.

Having a single affected child in a family with no previous history of a presumed genetic disorder might suggest the possibility of a recessive disorder, with a 1 in 4 risk that each subsequent child would be affected.

Alternatively, it could be a dominant disorder and the affected individual could be a heterozygote. In that case, one parent carries the disease gene but does not display the phenotype, or the disorder is due to a *de novo* mutation (see below).

One possible way to work out the mode of inheritance is to study multiple families with the same disorder and calculate the overall proportion of affected children (called the *segregation ratio)*. But there are many difficulties with this approach. First, the disorder may be heterogeneous and be due to different genes in different families. Secondly, the total numbers of children who can be studied are often too small to get reliable estimates.

There are also problems in how the families are ascertained (that is, in finding the people and families who will be studied). In the pre-genomics era trying to establish that a disorder was autosomal recessive was difficult. Then the priority would have been to collect a set of families and try to show a segregation ratio of 1 in 4. However, there was the complication of *ascertainment bias:* if there is no independent way of recognizing carriers, the families will be identified only through an affected child (families with two carrier parents and only unaffected children would seem perfectly normal and not be included).

Happily, in the genomics era underlying disease genes can quickly be found even for rare single-gene disorders. Rapid next-generation DNA sequencing is now being widely used to screen *exomes* (in practice, all exons of protein-coding genes) of affected individuals with the same condition. As is described in later chapters, genes underlying some rare single-gene disorders have been successfully identified after sequencing exomes from only a very few unrelated individuals with the disorder.

For a single-gene disorder, the observed incidence of mutant alleles in a defined population can be quite stable over time. A proportion of mutant alleles are transmitted from one generation to the next, and a proportion are lost because some individuals possessing mutant alleles do not transmit them. To keep the frequency of mutant alleles constant, new mutations

make up for the loss of mutant alleles that are not transmitted to the next generation.

Persons who have a severe disorder usually do not reproduce or have a much-reduced reproductive capacity (unless the disorder is not manifested until later in life). In severe autosomal recessive conditions, however, for each affected individual there are very many asymptomatic carriers who can transmit mutant alleles to the next generation. Because only a very small proportion of mutant alleles go untransmitted, the incidence of new mutation is low.

For severe dominant disorders, the mutant alleles are concentrated in affected individuals. If most individuals who carry the disease allele do not reproduce (because the disorder is congenital, say), the incidence of new mutation will be very high. If, however, there is a relatively late age at onset of symptoms, as with Huntington disease, individuals with the mutant allele may reproduce effectively, and the rate of new mutation may be very low.

For severe X-linked recessive disorders, the incidence of new mutation will also be quite high to balance the loss of mutant alleles when affected males do not reproduce. However, female carriers will usually be able to transmit mutant alleles to the next generation.

As a result of a new mutation, an affected person may be born in a family with no previous history of the disorder and would present as an isolated *(spo- radic)* case. In rare disorders that have not been well studied, a sporadic case poses difficulty for calculating the risk that subsequent children could also be affected. The affected individual could be a heterozygote (as a result of *de novo* mutation, or the failure of the disorder to be expressed in one parent), but alternatively could be a homozygote born to carrier parents, or a hemizygous boy whose mother is a carrier of an X-linked recessive condition.

Post-zygotic mutations and mosaicism

Most mutations arise as a result of endogenous errors in DNA replication and repair. Mutations can occur during gametogenesis and produce sperm and eggs with a new mutant allele. In addition, *de novo* pathogenic mutations can also occur at any time in post-zygotic life. As a result of post-zygotic mutations, each individual person is a genetic **mosaic** with genetically distinct populations of cells that have different mutational spectra.

Post-zygotic mutations may result in somatic mosaicism that will have consequences only for that individual (**Box 5.3**). But certain post-zygotic mutations, often occurring comparatively early in development, may also result in *germline mosaicism*. A person who has a substantial proportion of mutant germline cells (a germline mosaic or gonadal mosaic) may not show any symptoms but will produce some normal gametes and some mutant gametes. The risks of having a subsequently affected child are much higher than if an affected child carries a mutation that originated in a meiotic division.

BOX 5.3 POST-ZYGOTIC MUTATIONS AND WHY WE ARE ALL GENETIC MOSAICS

A pathogenic new mutation can be imagined to occur during gamete formation in an entirely normal person. Most mutations arise as a result of endogenous errors in DNA replication and repair, and although mutations do occur during gametogenesis and produce sperm and eggs with a new heritable mutant allele, they can also occur at any time in post-zygotic life. As a result of post-zygotic mutations, each individual person is a genetic *mosaic* with genetically distinct populations of cells that have different mutational spectra.

Human mutation rates are around 10^{-6} per gene per generation, and so a person with a wild-type allele at conception has a roughly one in a million chance of transmitting it to a child as an altered (mutant) allele. In this case we are considering the chance of a mutation occurring in a lineage of germline cells from zygote to gamete, involving a series of about 30 cell

divisions in females and several hundred divisions in males (about 400 by age 30 and increasing by about 23 per year because spermatogenesis continues through adult life—see Figure 7.5 on page 190).

Now consider post-zygotic mutations in somatic cell lineages. The journey from single-celled zygote to an adult human being involves a total of about 10^{14} mitotic cell divisions. With so many cell divisions, post-zygotic mutation is unavoidable—we must all be mosaics for many, many mutations. Having so many potentially harmful somatic mutations is usually not a concern because the number of cells that will fail to function correctly is normally very small. A cell will usually function normally after sustaining a harmful mutation in a gene that is not normally expressed in that cell type, and even if the cell does function abnormally as a result of mutation it might not give rise to many mutant descendants.

A person may be at risk of disease, however, if a mutated cell is able to give rise to substantial numbers of descendant cells that act abnormally (**Figure 1**). The biggest disease risk posed by post-zygotic mutations is that they set off or accelerate a process that leads to cancer. As we describe in Chapter 10, cancers are unusual in that although they can be inherited, the biggest contribution to disease comes from somatic mutations.

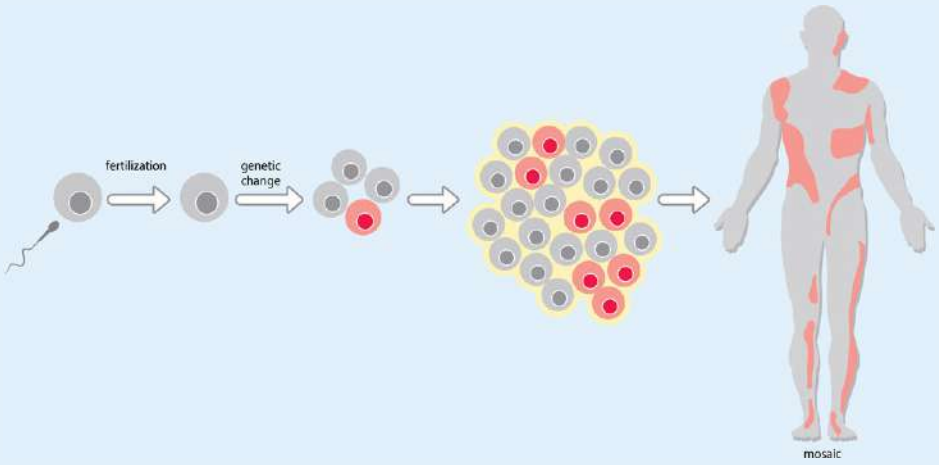


Figure 1 **Genetic mosaicism**. As illustrated here, post-zygotic mosaicism may often have consequences just for the individual who possesses the mutant cells; that is, the mutation affects somatic cells only. Sometimes, however, post-zygotic mutations can

occur in germ-cell precursors (*germline mosaicism*), and that has important implications concerning the possibility of transmitting a disorder.

Heterogeneity in the correspondence between phenotypes and the underlying genes and mutations

There is no one-to-one correspondence between phenotypes and genes. Three levels of heterogeneity are listed below. As we will see below and in later chapters, both nongenetic factors (environmental and epigenetic) and additional genetic factors can also influence the phenotypes of single-gene disorders.

Locus heterogeneity

The same clinical phenotype can often be produced by mutations in genes at two or more loci. The different genes often make related products that work together as a complex or in a common pathway; sometimes one gene is the primary regulator of another gene.

Locus heterogeneity explains how parents who are both affected with the same common recessive disorder produce multiple unaffected children. Recessively inherited deafness is the classic example (sensorineural hearing impairment mostly shows autosomal recessive inheritance, and deaf people often choose to have children with another deaf person). If two deaf parents are homozygous for mutations at the same gene locus, one would expect that all their children would also have impaired hearing. If, instead, the parents are homozygous for mutations at two different recessive deafness loci, all their children would be expected to be double heterozygotes and have normal hearing (**Figure 5.10**).

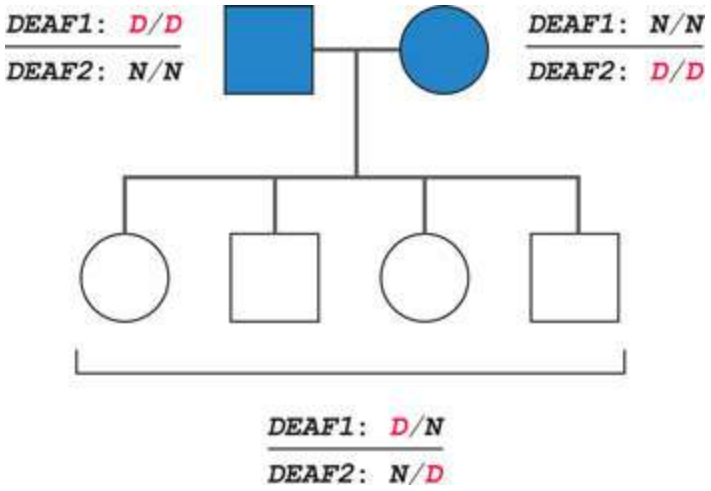


Figure 5.10 Locus heterogeneity explains why two parents with autosomal recessive deafness can consistently produce unaffected children. Imagine that the two parents are deaf because they have two mutant alleles at different autosomal recessive deafness loci, which we represent here as *DEAF1* and *DEAF2*. We represent normal alleles as *N* and deafness-associated alleles as D. In this case, sperm produced by the father would carry the *DEAF1*D* allele and the *DEAF2*N* allele, and eggs produced by the mother would carry the *DEAF1*N* allele and the *DEAF2*D* allele. All children would therefore be unaffected because they would be heterozygous at both loci. The normal phenotypes of each child result from complementation between normal alleles at the two loci. If, instead, both parents had autosomal recessive deafness caused by different mutations in the *same* gene, all their children would be expected to be deaf as a result of inheriting two mutant alleles at that locus.

As the underlying genes for single-gene disorders become known, it has become clear that very many conditions show locus heterogeneity. One might anticipate that many different genes contribute at different steps to broad general pathways (responsible for hearing or vision, for example). It is therefore unsurprising that autosomal recessive deafness or retinitis pigmentosa (hereditary retinal diseases with degeneration of rod and cone photoreceptors) can result from mutations in different genes.

More specific phenotypes can also be caused by mutations at any one of many different gene loci. Usher syndrome, for example, involves profound sensorineural hearing loss, vestibular dysfunction, and retinitis pigmentosa;

autosomal recessive forms can be caused by mutations at any one of at least 11 different gene loci.

Bardet-Biedl syndrome (PMID 20301537) provides another illustrative example. It is a *pleiotropic* disorder (many different body systems and functions are impaired) and the primary features are: degeneration of light-sensitive cells in the outer regions of the retina (causing night blindness, tunnel vision, reduced visual acuity), learning disabilities, kidney disease, extra toes and/or fingers, obesity, and abnormalities of the gonads. Autosomal recessive inheritance is the typical inheritance pattern, and the disorder is caused by mutations in any of at least 21 genes, all involved in regulating how cilia function (**Figure 5.11**).

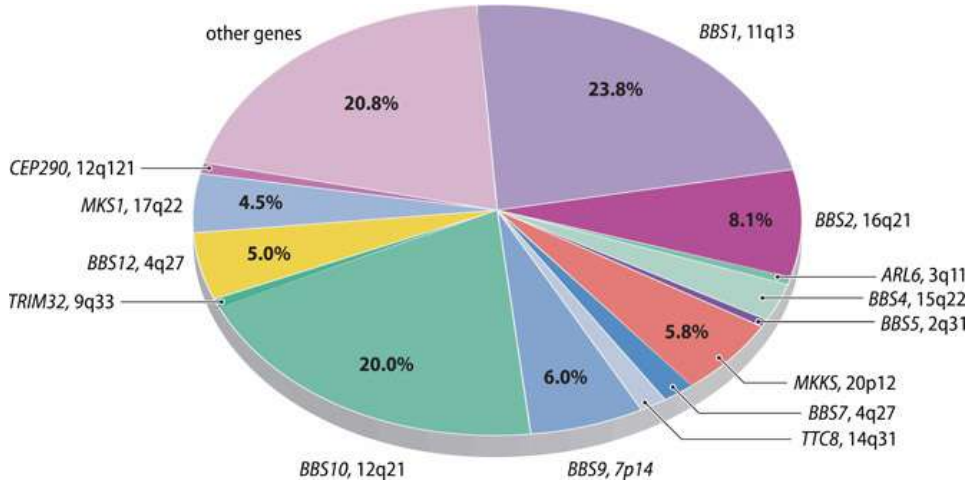


Figure 5.11 Extraordinary locus heterogeneity for Bardet-Biedl syndrome (BBS). Segments represent the proportion of total mutant alleles attributable to the first 14 genes known to be mutated in BBS, with 20.8 % initially unidentified genes represented here by the segment labeled "Other genes". Seven of the genes – *BBS1, BBS2, MKKS, BBS9, BBS10, BBS12,* and *MKS1* – account for ~70 % of the identified pathogenic mutations in BBS. Note: some of the genes are also mutated in other disorders, such as *MKS1* (in Meckel syndrome and Joubert syndrome) and *MKKS* (in McKusick-Kaufman syndrome). Since this figure was first published, seven more identified BBS genes have been described (see PMID 29487844). (Adapted from Zaghloul NA & Katsanis N [2009] *J Clin Invest* 119:428–437; PMID 19252258. With permission from the American Society for Clinical Investigation.)

Allelic and phenotypic heterogeneity

Many different mutations in one gene can have the same effect and produce similar phenotypes. For example, b-thalassemia results from a deficiency of b-globin and can arise by any number of different inactivating mutations in the hemoglobin b chain *(HBB)* gene. Different mutations in a single gene can also often result in different phenotypes. That can arise in two ways: either different types of mutation somehow have different effects on how the underlying gene works—which we consider here—or some factors outside the disease locus have varying effects on the phenotype (described later).

Phenotype variation due to different mutations at a single gene locus may differ in degree (severe or mild versions of the same basic phenotype) or be extensive and result in rather different disorders. For example, Duchenne and Becker muscular dystrophies (OMIM 310200 and 300376, respectively) represent severe and mild forms of the same type of muscular dystrophy and are both examples of dystrophinopathies (PMID 20301298). More extreme phenotype heterogeneity can result from mutations at some genes (see the example of the lamin A/C gene in **Table 5.1**).

TABLE 5.1 REMARKABLE HETEROGENEITY OF CLINICAL PHENOTYPES RESULTING FROM MUTATION IN THE LAMIN A *(LMNA)* GENE

Class of disorder	Disorder	Inheritance pattern	OMIM No.
Lipodystrophy	lipodystrophy, familial partial, type 2	AD	151660
	mandibulosacral dysplasia type A with lipodystrophy	AR	248370
	Emery-Dreifuss muscular dystrophy type 2	AD	181350
	Emery-Dreifuss muscular dystrophy type 3	AR	181350

Class of disorder	Disorder	Inheritance pattern	OMIM No.
	congenital muscular dystrophy	AD	613205
	cardiomyopathy, dilated type IA	AD	115200
	Malouf syndrome (cardiomyopathy, dilated, with hypertrophic hypogonadism)	AR	212112
	heart-hand syndrome, Slovenian type	AD	610140
Neuropathy	Charcot-Marie-Tooth disease, type 2B1	AR	605588
Progeria	Hutchinson-Gilford progeria syndrome	AD, AR	176670

Clinical phenotypes can also vary between affected members of the same family even although they have identical mutations. As we saw in Section 5.2, heteroplasmy can explain divergent phenotypes in family members affected by a mitochondrial DNA disorder. But single-gene disorders can also show intrafamilial variation in phenotype that may be due to genetic and nongenetic factors as described below.

Nonpenetrance and age-related penetrance

The **penetrance** of a single-gene disorder is the probability that a person who has a mutant allele will express the disease phenotype. Dominantly inherited disorders, by definition, are manifested in heterozygotes and might be expected to show 100 % penetrance. That might be true for certain dominant disorders. For many others, however, penetrance is more variable and the disorder can sometimes appear to skip a generation so that a person who must have inherited the disease allele is unaffected (**nonpenetrance**—see **Figure 5.12**).

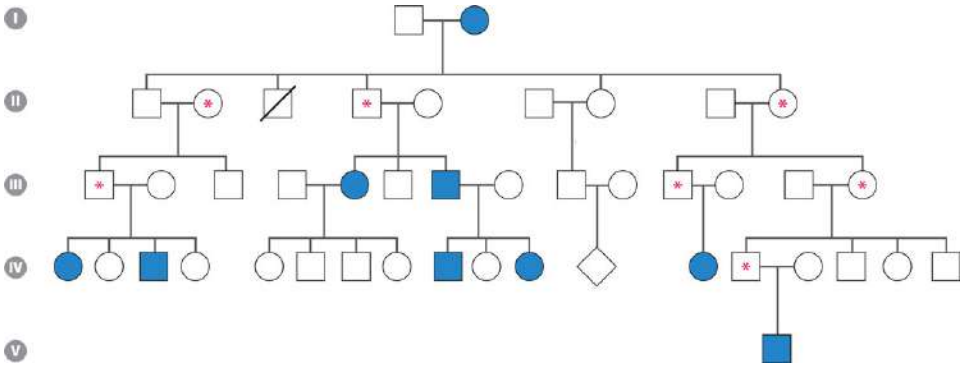


Figure 5.12 Nonpenetrance in an autosomal dominant disorder. Individuals with a red asterisk are asymptomatic disease gene carriers: they have inherited a mutant allele ultimately from the affected great-great-grandmother in generation I, but none of them expresses the disease phenotype. In this example, the disorder is evident only in individuals who have inherited a mutant allele from their father (in each case the unaffected individuals with a red asterisk inherited a mutant allele from their mother). As described in the text, an epigenetic mechanism known as *imprinting* can result in this type of parent-of-origin effect on the phenotype.

Nonpenetrance should not be viewed as surprising. Even in single-gene disorders—in which, by definition, the phenotype is largely dictated by the genotype at just one locus—other genes can play a part, as can epigenetic and environmental factors.

Variable age at onset in late-onset disorders

A disease phenotype may take time to manifest itself. If a disorder is present at birth, it is said to be congenital. In some disorders, however, there is a late age at onset so that the penetrance is initially very low but then increases with age. Age-related penetrance means a late onset of symptoms, and quite often the disease first manifests in adults.

The slow development of disease in adult-onset disorders may occur in different ways. Harmful products may be produced slowly but build up over time, for example. If pathogenesis involves a gradual process of cell death, it may take some time before the number of surviving cells drops to

critically low levels that produce clinical symptoms. In hereditary cancers, a mutation is inherited at a tumor-suppressing gene locus and a second, somatic mutation is required to initiate tumor formation. The second mutation occurs randomly, but the probability of a second mutation increases with time and therefore with age.

Huntington disease is a classic example of a late-onset single-gene disorder. In this case, mutant alleles produce an abnormal protein that is harmful to cells and especially toxic to neurons. The loss of neurons is gradual but eventually results in a devastating neurodegenerative condition. Huntington disease is highly penetrant. The onset of symptoms typically occurs in middle to late adult life, but juvenile forms are also known (Figure 5.13).

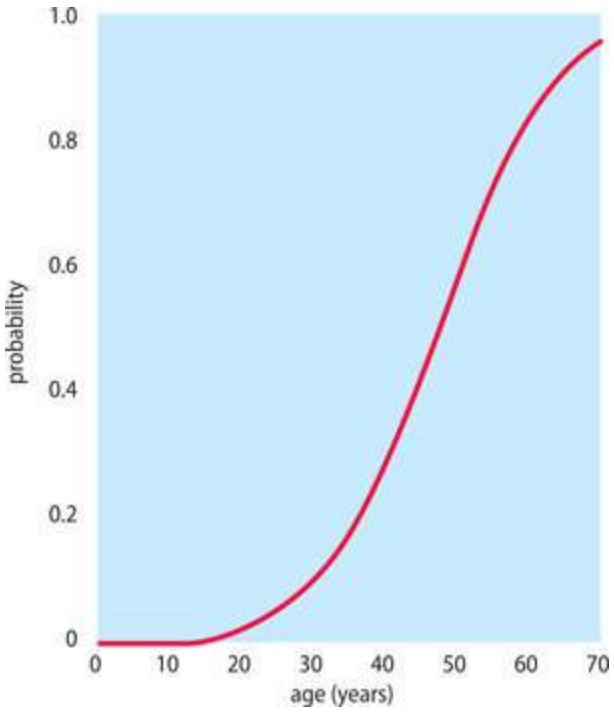


Figure 5.13 Age-related onset of Huntington disease. The curve shows the probability that an individual carrying a Huntington disease allele will have developed symptoms by a given age. (From Harper PS [2010] *Practical Genetic Counselling*, 7th ed. With permission from Taylor & Francis Group LLC.)

Age-at-onset curves for late-onset disorders are used in genetic counseling to calculate the chance that an asymptomatic person at risk of

developing the disease carries the mutation. In Huntington disease an unaffected person who has an affected parent will have a 50 % *a priori* risk that decreases with age (see Figure 5.13); if one is still free of symptoms by age 60, for example, the chance of developing the disease falls to less than 20 %.

Phenotypes resulting from mutation in mitochondrial DNA are highly variable because of the special mitochondrial property of heteroplasmy (see Section 5.2). Some types of Mendelian disorders, notably dominant phenotypes, are also prone to variable expression, and different family members show different features of disease (sometimes called *variable expressivity*—see **Figure 5.14** for an example pedigree). But, like nonpenetrance, variable phenotype expression is occasionally seen in recessive pedigrees.

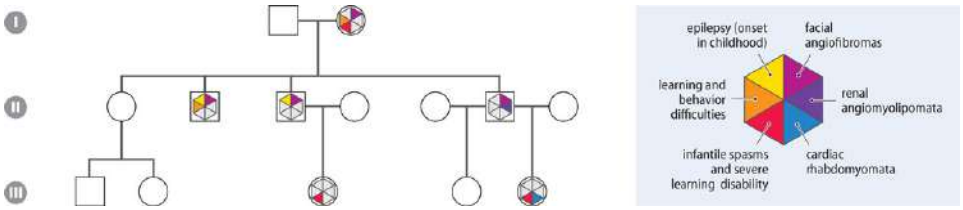


Figure 5.14 Variable phenotypes in a tuberous sclerosis family. Tuberous sclerosis is an autosomal dominant disorder caused by mutations in either the *TSC1* or *TSC2* gene. These two genes make two subunits of a tumor suppressor protein complex that regulates cell growth and proliferation. The disorder affects multiple body systems with characteristic tumor-like lesions in the brain, skin, and other organs, and is often associated with seizures and learning difficulties. However, as is evident in this family from the northeast of England, there can be considerable differences in *expressivity* of the disorder. (Pedigree information provided by Dr Miranda Splitt, Northern Region Genetics Service UK.)

Nonpenetrance can be regarded as an extreme endpoint of variable expression, and the factors that produce variable expression of phenotypes within families are the same as those that result in nonpenetrance. They include nongenetic factors—epigenetic regulation and environmental factors (**Figure 5.15B**) and also stochastic factors. Additional genetic

factors are also involved, notably *modifier genes* that regulate or interact with a Mendelian locus, affecting how it is expressed. Different alleles at a modifier gene locus may have rather different influences on the expression of the Mendelian locus (Figure 5.15B).

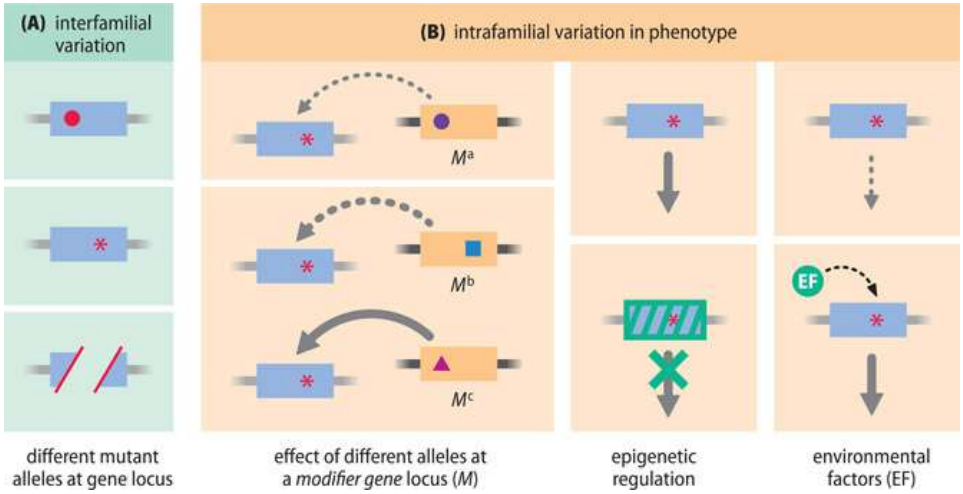


Figure 5.15 Main explanations for phenotype variation in Mendelian disorders. (A) *Interfamilial variation in phenotype*. Unrelated individuals with the same Mendelian disorder may often have different mutations (red symbols) at the disease gene locus with different consequences for gene expression and disease. (B) *Intrafamilial variation in phenotype*. Affected members of a single family can be expected to have the same mutation at the disease gene locus but can nevertheless show differences in phenotype because of genetic or nongenetic factors. In the former case, the affected individuals may have different alleles at one or more modifier gene loci. Modifier genes make products that interact with the primary gene locus so as to modulate the phenotype, and different alleles of a modifier gene can have different effects. Alternatively, nongenetic factors can explain phenotype variation; an example is epigenetic regulation, in which the disease allele can be differently regulated in some individuals by an altered chromatin conformation (green hatched box) or by variable exposure to an environmental factor (green circle) such as a specific virus or chemical during development *in utero*.

Imprinting

Certain phenotypes show autosomal dominant inheritance with parent-of-origin effects. Both sexes are affected, and the mutant allele can be transmitted by either sex but is expressed only when inherited from a parent of one particular sex. For some conditions, a mutant allele must be inherited from the father for the disease to be expressed (see Figure 5.12 for an example). For other conditions, such as Beckwith-Wiedemann syndrome, the disease phenotype is expressed only if the disease allele is inherited from the mother (**Figure 5.16**).

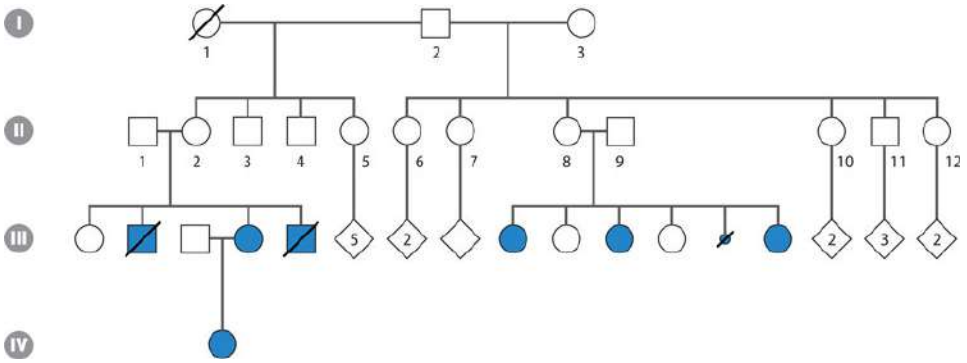


Figure 5.16 Parent-of-origin effect on the expression of an inherited disorder. This pedigree shows autosomal dominant Beckwith-Wiedemann syndrome (PMID 20301568), which manifests only when the underlying mutant allele is maternally inherited. The affected individuals in generation III must have inherited the mutant allele from their common grandfather I-2 but none of his 10 children in generation II have symptoms of disease, including two daughters, II-2 and II-8, who have gone on to have multiple affected children. (From Viljoen D & Ramesar R [1992] *J Med Genet* 29:221–225; PMID 1583639. With permission from BMJ Publishing Group Ltd.)

The parent-of-origin effects are due to an epigenetic mechanism known as imprinting, which we describe in detail in Chapter 6. The mutant allele that is not expressed is often described as the imprinted allele. Accordingly, Beckwith-Wiedemann syndrome is said to be paternally imprinted, because paternally inherited alleles are not expressed.

Anticipation

Some disorders show consistent generational differences in phenotype. Disorders such as fragile X mental retardation syndrome, myotonic dystrophy, and Huntington disease are caused by unstable mutations (often called *dynamic mutations)* whose characteristics can change after they undergo DNA replication. As a result, the phenotype can vary between affected individuals in families but in a directional way; that is, it can be expressed at an earlier age and become increasingly severe with each new generation of affected individuals. This phenomenon is known as **anticipation** (Figure 5.17). We consider the molecular mechanisms in detail in Chapter 7.

Figure 5.17 A three-generation family affected with myotonic dystrophy.
The degree of severity increases in each generation. The grandmother (right) is only slightly affected, but the mother (left) has a characteristic narrow face and somewhat limited facial expression. The baby is more severely affected and has the facial features of children with neonatal-onset myotonic dystrophy, including an open, triangular mouth. The infant has more than 1000 copies of the trinucleotide repeat, whereas the

mother and grandmother each have about 100 repeats. (From Jorde LB, Carey JC & Bamshad MJ [2009] *Medical Genetics*, 4th ed. With permission from Elsevier.)

5.4 ALLELE FREQUENCIES IN POPULATIONS

Genetic disorders that are comparatively common and serious have somehow avoided being eliminated by natural selection. This raises two questions. First, are high mutation rates enough to explain why harmful disease alleles persist? And if so, why should some single-gene disorders be comparatively common but others very rare? In this section we are concerned primarily with allele frequencies and the factors that affect them.

The frequency of a single-gene disorder in a population relates to the frequency in the population of pathogenic alleles at the relevant disease locus (or loci). A high disease allele frequency might result if a gene were to be particularly susceptible to mutation. Large genes may contain many repetitive sequences that confer structural instability, such as the very large dystrophin gene that is very prone to intragenic deletions and duplications.

Some of the most common single-gene disorders, such as sickle-cell anemia and the thalassemias, result from mutation in tiny genes—as we will see below, autosomal recessive disorders do not require high mutation rates to be common. Even in some autosomal dominant disorders, a high incidence of the disorder may not necessarily mean that the underlying gene loci have high mutation rates, as described below.

Some disorders may be caused by a *selfish mutation*. Achondroplasia (PMID 20301331) is a common single-gene disorder but is caused exclusively by mutation at just a single nucleotide, producing a highly specific change (glycineto-arginine substitution at residue 380) in the FGFR3 (fibroblast growth factor receptor type 3) protein. The nucleotide that is altered is not thought to be highly mutable. Instead, the mutation may promote its own transmission: male germ-line cells that contain it may have a proliferative advantage and make a disproportionate contribution to

sperm. As a result, there is a high allele frequency even although the mutation rate is not so exceptional. We consider selfish mutations in detail in Section 7.2.

We also need to explain why some single-gene disorders are common in some human populations but very rare in others. Cystic fibrosis is particularly common in northern European populations, for example, and sickle-cell anemia is especially frequent in tropical Africa but virtually absent from many other human populations.

In all of these considerations, what do we mean by a human *population?* We could mean anything from a small tribe to the whole of humanity. An idealized population would be large with no barriers to random mating; as we will see below, some important principles in population genetics are based on this kind of population.

In practice, mating is often far from random because of different types of barriers. Geographic barriers can mean that people who live in locations that are remote (or otherwise difficult to access) form populations with limited genetic diversity and with distinctive allele frequencies. But even within single cities there are also many ethnic populations with distinctive allele frequencies. And, as we will see below, even within these populations, mating is not random.

Allele frequencies and the Hardy-Weinberg law

The frequency of an allele in a population can vary widely from one population to another. The concept of the **gene pool** (all of the alleles at a specific gene locus within the population) provides the reference point for calculating allele frequencies (which are often inaccurately represented in the literature as *gene frequencies)*.

For a specific allele, say allele *A*1* at locus A, the **allele frequency** is the proportion of all the alleles in the population at locus *A* that are *A*1* and is given as a number between 0 and 1. Effectively, the allele frequency for *A*1* is the *probabil- ity* that an allele, picked at random from the gene pool, would be *A*l*.

The Hardy-Weinberg law

The Hardy-Weinberg law (or equilibrium, principle, theorem) provides a mathematical relationship between allele frequencies and genotype frequencies in an *idealized* large population where matings are random and allele frequencies remain constant over time.

Imagine that locus *A* has only two alleles, *A*1* and *A*2*, and that their respective frequencies are p and q (so that $p + q = 1$). The respective genotypes are combinations of two alleles at a time. To calculate the frequency of a genotype, we therefore first need to estimate the probabilities of picking first one specified allele from the gene pool (as the paternal allele, say), and then picking a second allele to be the maternal allele.

Imagine we pick *A*1* first (with a probability of p) and then we pick *A*1* again (with a probability of p). If the population is large, the two probabilities are independent events and so the joint probability of picking *A*1* first and then *A*1* again is the product of the two probabilities, namely p^2. This is the only way that we can arrive at the genotype *A*1.A*1,* whose frequency is therefore p^2 (**Figure 5.18**).

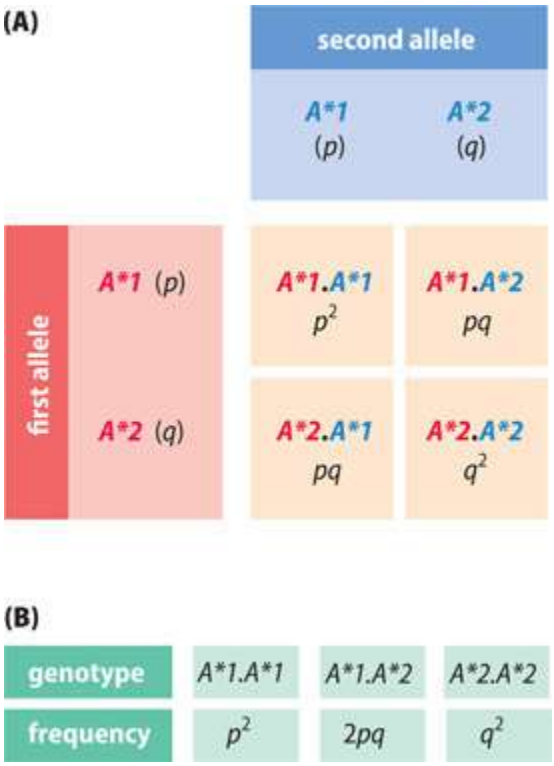


Figure 5.18 Visualizing how genotype frequencies are related to allele frequencies. In this example we consider a locus *A* that has two alleles, *A*1* and *A*2,* with respective frequencies *p* and *q*. Genotypes are unique combinations of *two* alleles, one from a father and one from a mother. (A) We can first construct a matrix of all possible pairwise allele combinations, whose frequencies are simply the products of the frequencies of the two alleles. (B) We then integrate any pairwise combinations that have the same two alleles *(A*1.A*2* is effectively the same as *A*2.A*1)* to get the frequencies of the three unique genotypes. Note that the Hardy-Weinberg law relates genotype frequencies to allele frequencies by a binomial expansion: $(p + q)^2 = p^2 + 2pq + q^2$ for two alleles (as shown here), $(p + q + r)^2$ for three alleles, $(p + q + r + s)^2$ for four alleles, and so on.

Now consider the genotype *A*1.A*2*. We can get this in two ways. One way is to first pick *A*1* (probability *p*) and then *A*2* (probability *q*), giving a joint probability of *pq*. But a second way is to pick *A*2* first (probability *q*) and then pick A*1 (probability *p*), again giving a combined probability of *pq*. As a result, the frequency of the genotype *A*1.A*2* is 2*pq*.

In summary, in a suitably ideal population, the Hardy-Weinberg law gives the frequencies of homozygous genotypes as the square of the allele frequency, and the frequencies of heterozygous genotypes as twice the product of the two allele frequencies. An important consequence is that if allele frequencies in a population remain constant from generation to generation, the genotype frequencies will also not change.

Applications and limitations of the Hardy-Weinberg law

The major clinical application of the Hardy-Weinberg law is as a tool for calculating genetic risk. In a family with a single-gene disorder, only one or two mutant alleles are normally found in the causative gene, but within a population there may be many different mutant alleles at the disease locus. To apply the Hardy-Weinberg law to single-gene disorders, all the different mutant alleles are typically lumped together to make one disease allele. That is, we envisage just two alleles according to their effect on the disease phenotype: a normal allele (*N*), with no effect on the phenotype, and a disease allele *(D),* which can be *any* mutant allele. If we assign frequencies of p for allele *N* and q for allele *D*, the genotype frequencies would be as follows: p^2 for *NN* (normal homozygotes), $2pq$ for *ND* (heterozygotes), and q^2 for *DD* (disease homozygotes).

Practical application of the Hardy-Weinberg law to single-gene disorders is largely focused on autosomal recessive disorders, where it allows the frequency of carriers to be calculated without having to perform relevant DNA tests on a large number of people (**Box 5.4**). Its utility depends on certain assumptions—notably random mating and constant allele frequencies—that may not be strictly upheld. As described below, allele frequencies can change in populations, but the changes are often slow and in small increments, and often have minor effects in disturbing the Hardy-Weinberg distribution of genotypes. However, certain types of nonrandom mating can substantially upset the relative frequency of genotypes predicted by the Hardy-Weinberg law.

BOX 5.4 USING THE HARDY-WEINBERG LAW TO CALCULATE CARRIER RISKS FOR AUTOSOMAL RECESSIVE DISORDERS

Genetic counseling for autosomal recessive conditions often requires calculations to assess the risk of being a carrier. The proband who seeks genetic counseling is typically a prospective parent with a close relative who is affected. He/she is worried about the high risk of being a carrier and then about the risk that his/her spouse could also be a carrier.

The proband's chance of being a carrier can be calculated by using the principles of Mendelian inheritance, but the Hardy-Weinberg law is used to calculate the risk that his/her spouse could also be a carrier. If both parents were to be carriers, each child would have a 1 in 4 risk of being affected.

Take the specific example in Figure 1. The healthy proband (arrowed) has a sister with cystic fibrosis and is worried about the prospect that he and his wife might have a child with cystic fibrosis. His wife is Irish, and the Irish population has the highest incidence of cystic fibrosis in the world, affecting one birth in 1350.

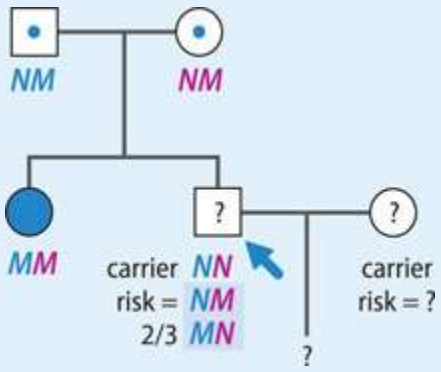


Figure 1 Using a combination of Mendelian principles and the Hardy-Weinberg law to estimate disease risk. The arrow indicates the proband. *N*, normal allele. *M*, mutant allele.

The proband's parents can be presumed to be carriers, each with one normal allele *N* and one mutant allele M. Because the proband is healthy,

he must have inherited one of three possible combinations of parental alleles: *N* from both parents (homozygous normal); *N* from father and *M* from mother (carrier); and *M* from father and *N* from mother (carrier). So from Mendelian principles, he has a risk of 2/3 of being a carrier (see Figure 1).

The risk that his wife is a carrier is the same as the probability that a person, picked at random from the Irish population, is a carrier. If we assign a frequency of p for the normal allele and q for the cystic fibrosis allele, the Hardy-Weinberg law states that the frequency of affected individuals will be q^2 and the frequency of carriers will be 2pq. Because population surveys show that cystic fibrosis affects 1 in 1350 births in the Irish population, $q^2 = 1/1350$ and so $q = 1/\sqrt{1350}$, or $1/36.74 = 0.027$.

Since $p + q = 1$, the value of $p = 0.973$. The risk of the wife being a carrier ($2pq$) is therefore $2 \times 0.973 \times 0.027 = 0.0525$, or 5.25 %. The combined risk that both the proband and his wife are carriers is $2/3 \times 0.0525 = 0.035$, or 3.5 %.

For rare autosomal recessive disorders, the value of p very closely approximates 1, so the carrier frequency can be taken to be $2q$. However, if the disorder is especially rare, the chances that the prospective parents are consanguineous is much higher, making the application of the Hardy-Weinberg law much less secure.

Nonrandom mating

In addition to geographical barriers to random mating, people also preferentially select mates who are similar to themselves in different ways. They may be members of the same ethnic group and/or sect, for example. Because breeding is less frequent between members from different communities, allele frequencies can vary significantly in the different communities. Geneticists therefore need to define populations carefully and calculate genetic risk by using the most appropriate allele frequencies.

Additional types of assortative mating occur. We also tend to choose a mate of similar relative stature and intelligence to us, for example. Positive assortment mating of this type leads to an increased frequency of homozygous genotypes and a decreased frequency of heterozygous genotypes. It extends to medical conditions. People who were born deaf or blind have a tendency to choose a mate who is similarly affected.

Inbreeding is a powerful expression of assortative mating that is quite frequent in certain societies and can result in genotype frequencies that differ significantly from Hardy-Weinberg predictions. Consanguineous mating results in an increased frequency of mating between carriers and a correspondingly increased frequency of autosomal recessive disease.

Ways in which allele frequencies change in populations

Allele frequencies can change from one generation to the next in different ways. Often changes in allele frequency are quite slow, but occasionally the composition of populations can change quickly, producing major shifts in allele frequency. Principal ways in which alleles change in the frequency of a population are listed below.

- *Purifying selection*. If a person affected by genetic disease is unlikely to reproduce, disease alleles are lost from the population (a form of negative natural selection). This effect is much more pronounced in early-onset dominant conditions, in which—with the exception of nonpenetrance—anyone with a mutant allele is affected by the time of puberty.
- *New mutations*. New alleles are constantly being created by the mutation of existing alleles. Some mutations produce new disease alleles by causing genes to lose their function or to function abnormally. There are numerous different ways in which a "forward" mutation can cause a gene to lose its function, but a "back mutation" *(revertant* mutation) that can restore the function of a nonfunctioning allele has to be very specific and so is comparatively very rare.

- *Influx of migrants*. If a population absorbs a large influx of migrants with rather different allele frequencies, then the overall gene pool will change.
- *Random sampling of gametes*. Only a certain proportion of individuals within a population reproduce. Out of all the alleles within the population, therefore, only those present in people who reproduce can be transmitted to the next generation. That is, a *sample* of the total alleles in the population is passed on and that sample is never exactly representative of the total population for purely statistical reasons. The smaller the size of a population, the larger will be the random fluctuations in allele frequency. This effect is known as genetic drift and in small populations it can cause comparatively rapid changes in allele frequencies between generations (**Figure 5.19**).

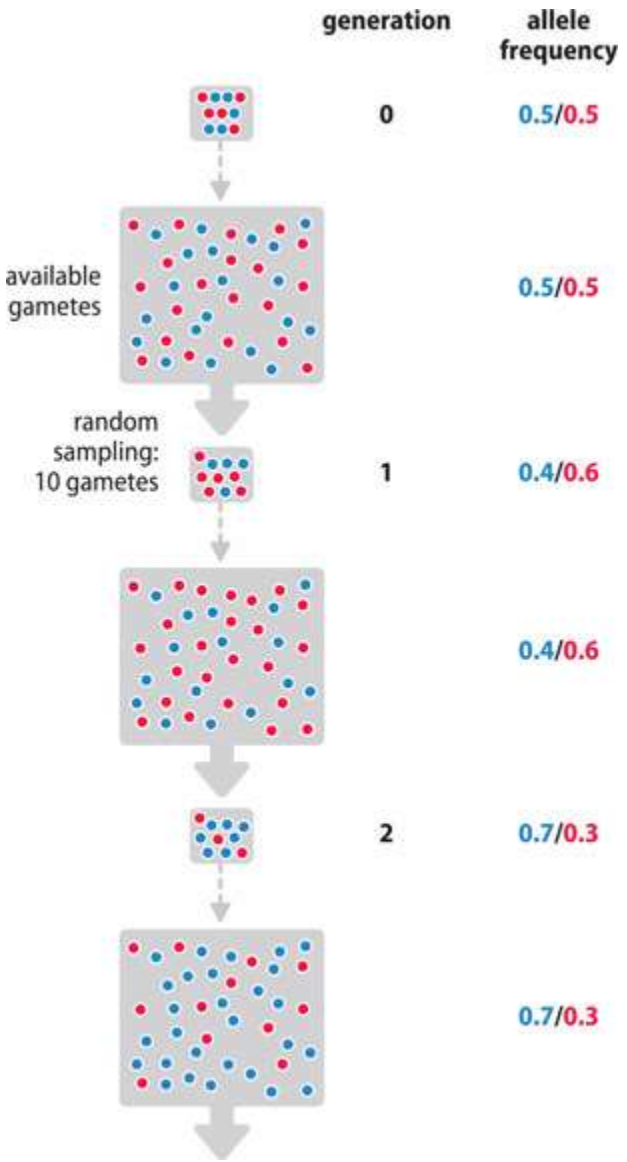


Figure 5.19 Random sampling of gametes in small populations can lead to considerable changes in allele frequencies. Small boxes represent gametes transmitted by reproducers to the next generation; large boxes represent all available gametes in the population. The comparative frequencies of the red and blue alleles can change significantly between generations when by chance the samples of transmitted gametes have allele frequencies that are rather different from the allele frequencies in the population. Such *genetic drift* is significant in small populations. (Adapted from Bodmer WF and Cavalli-Sforza LL [1976] *Genetics, Evolution and Man*. With permission from WH Freeman & Company.)

Population bottlenecks and founder effects

Genetic drift is most significant when population sizes are small. There have been several occasions during our evolution when the human population underwent a *population bottleneck,* a severe reduction in size before the reduced population (now with much less genetic variation) expanded again (Figure 5.20A). As a result, genetic variation in humans is very much less than in our nearest relative, the chimpanzee.

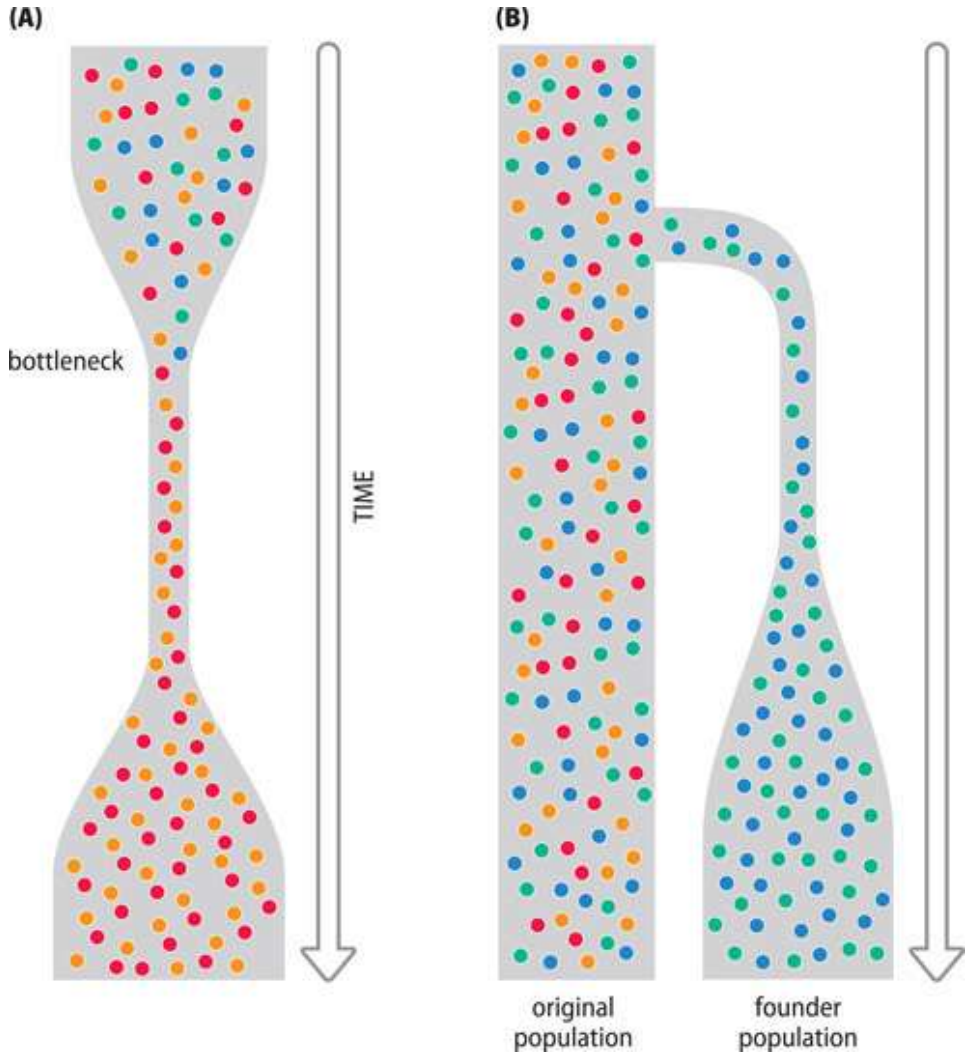


Figure 5.20 Altered allele frequencies after a population bottleneck and formation of a founder population. (A) In a population bottleneck, a severe reduction in the size of the population can lead to altered allele frequencies and much less genetic variation in the subset of the surviving population. Subsequent expansion will reestablish a large

population but with reduced genetic variation compared with the time before the bottleneck. (B) A small group of individuals, with a subset of the genetic variation of the larger population, migrate to establish a separate colony (founder population) that can expand but continues to show different allele frequencies from the original foundation. In both images the vertical arrows indicate the passage of time.

Another type of population reduction has periodically happened during human migrations, when a small group of individuals emigrated to form a separate colony. Again, the small population would represent a subset of the genetic variation in the original population and have different allele frequencies. Subsequent expansion of the founding colony would lead to a new population that continued to have limited genetic variation and distinctive allele frequencies reflecting those of the original settlers (a **founder effect**—see Figure 5.20B).

If a founder colony happens to have an increased frequency of a disease allele, the new population that will descend from it can be expected to have an increased frequency of the disease. Various populations throughout the world have elevated frequencies of certain single-gene disorders as a result of a founder effect. In autosomal recessive disorders, the great majority of mutant alleles are found in asymptomatic carriers who transmit the mutant alleles to the next generation.

The Finnish and Ashkenazi Jewish populations have been particularly amenable to investigations of founder effects because of rapid recent population expansion, high education levels, and very well-developed medical services. After the introduction of agriculture from the Middle East in prehistoric times, Finland was one of the last regions of Europe to be populated, and the major expansion that led to the present population began only 2000–2500 years ago as migrants entered southern Finland. Thereafter, in the seventeenth century a second large population expansion began with the occupation of the uninhabited north of Finland.

Ashkenazi Jews (descended from a population that migrated to the Rhineland in the ninth century and from there to different countries in eastern Europe) and Sephardic Jews (primarily from Spain, Portugal, and

north Africa) have been distinct populations for more than a thousand years. Until just a few hundred years ago, Ashkenazi Jews used to represent a minority of Jews, but they have undergone a rapid population expansion and now account for 80 % of the global Jewish population. Founder effects have been documented in many other populations; see **Table 5.2** for examples.

TABLE 5.2 EXAMPLES OF SINGLE-GENE DISORDERS THAT ARE COMMON IN CERTAIN POPULATIONS BECAUSE OF A FOUNDER EFFECT

Disorder and inheritance (OMIM)	Population	Comments
Aspartylglucosaminuria; AR (208400)	Finnish	carrier frequency = 1 in 30
Ellis-van Creveld syndrome; AR (225500)	Amish, Pennsylvania	carrier frequency ≈ 1 in 8. Traced to a single couple who immigrated to Pennsylvania in 1774
Familial dysautonomia; AR (223900)	Ashkenazi Jews	carrier frequency = 1 in 30
Hermansky-Pudlak syndrome; AR (203300)	Puerto Ricans	thought to have been introduced by migrants from southern Spain
Alzheimer disease type 3, early onset; AR (607822)	in remote villages in the Andes	all descended from a couple of Basque origin who settled in Colombia in the early 1700s
Huntington disease (HD); AD (143100)	in fishing villages around Lake Maracaibo, Venezuela	more people with HD here than in rest of world. About 200 years ago, a single woman with the HD allele bore 10 children. Many current residents of Lake Maracaibo can trace their ancestry and disease-causing allele back to this lineage

Disorder and inheritance (OMIM)	Population	Comments
Myotonic dystrophy, type I, AD (160900)	in Saguenay-Lac-Saint-Jean, Quebec	prevalence of 1 in 500 (30–60 times more frequent than in most other populations). Introduced by French settlers

A distinguishing feature of a founder effect is that affected individuals will usually have mutant alleles with the same ancestral mutation. For example, affected individuals in nine Amish families with Ellis-van Creveld syndrome were shown to be homozygous for the same pathogenic mutation in the *EVC* gene and for a neighboring nonpathogenic sequence change that is absent from normal chromosomes. In this case, genealogy studies were able to confirm a founder effect: all affected individuals could trace their ancestry to the same couple, a Mr Samuel King and his wife, who immigrated in 1774.

Mutation versus selection in determining allele frequencies

If we consider stable, large populations (so that migrant influx and genetic drift are not significant factors), the frequencies of mutant alleles (and genetic diseases) in a population are determined by the balance between two opposing forces: mutation and selection.

Purifying selection removes disease alleles from the population when a disorder causes affected individuals to reproduce less effectively than the normal population. The genetic term **fitness** (***f***) is applied here and is really a measure of reproductive success: it uses a scale from 0 to 1 to rank the capacity of individuals to reproduce and have children who survive to a reproductive age. Thus, a fitness of 0 (genetic lethal) means consistent failure to reproduce, and so mutant alleles are not transmitted vertically to descendants. Loss of mutant alleles from the population by purifying selection is balanced by the creation of new mutant alleles by fresh mutation, keeping constant the disease allele frequency in the population.

For autosomal dominant disorders, all people who have a disease allele might be expected to be affected (if we discount nonpenetrance). Yet, according to the disorder, the fitness of individuals varies enormously. In many cases, affected individuals have severely or significantly reduced fitness. However, individuals affected by a late-onset disorder can have fitness scores that approach those of normal individuals—they are healthy in their youth and can reproduce normally (Figure 5.21 gives some examples).

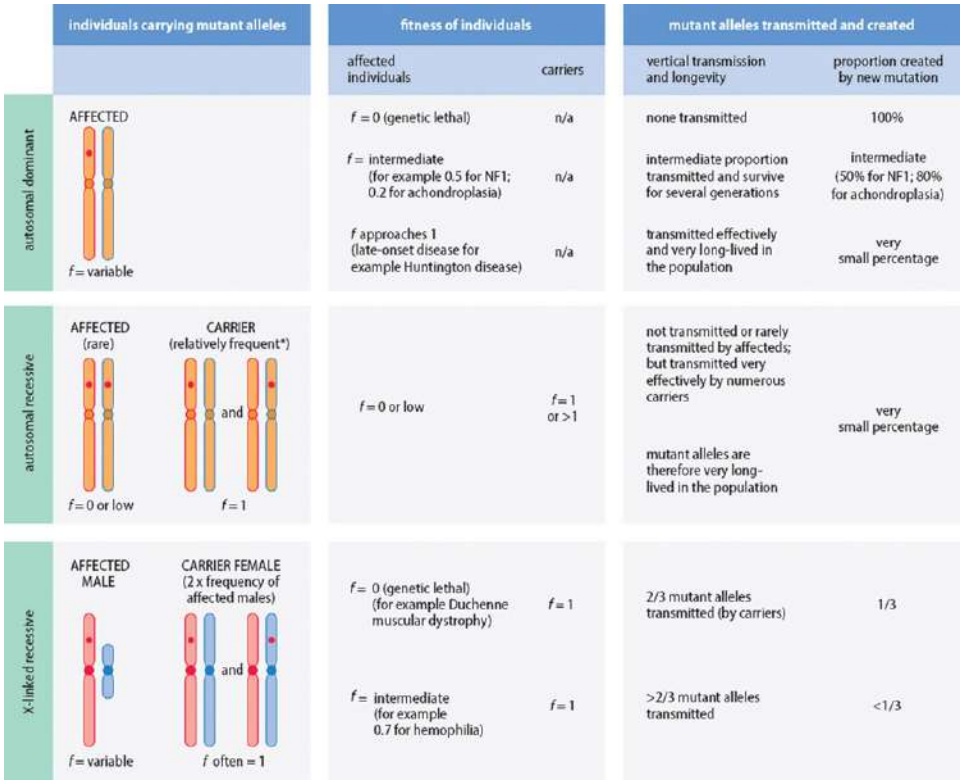


Figure 5.21 Fitness of individuals and mutant allele transmission/creation in single-gene disorders. Note that carriers of certain autosomal recessive disorders may have a higher fitness than normal individuals *(heterozygote advantage*—see following page).

For recessive disorders, mutant alleles are also found in carriers who have a single mutant allele. In autosomal recessive disorders, carriers vastly outnumber affected individuals. Recall the Hardy-Weinberg law that gives a ratio of $2pq$ (carriers) to q^2 (affecteds) = $2p/q \approx 2/q$ (p is very close to 1 for

almost all recessive disorders). To take one example, cystic fibrosis occurs in roughly 1 in 2000 births in northern European populations, so $q^2 = 1/2000$. This gives $q \approx 1/45$ and $2/q \approx 2/(1/45) = 90$. That is, there would be about 90 carriers of cystic fibrosis for each affected individual in this population. Because carriers of autosomal recessive disease are normally asymptomatic, they are very effective at transmitting mutant alleles and so new mutations are rare in autosomal recessive disease.

For X-linked recessive disorders, there are two female carriers per affected male. This happens because mutant alleles that reside on an X chromosome get transferred by recombination between three types of X chromosome: a single X chromosome in males and two X chromosomes in females. If we discount manifesting heterozygotes and take an approximation that the fitness of carriers is close to 1, then for conditions in which affected males do not reproduce, natural selection removes 1 out of 3 mutant alleles from the population. Because the lost alleles are replaced by new mutant alleles, 1 out of 3 mutations are new mutations.

Heterozygote advantage: when natural selection favors carriers of recessive disease

We saw above how some populations have a particularly high incidence of a genetic disorder as a result of a founder effect. Another reason why a recessive condition may be especially common in one population is that under certain conditions a type of natural selection can favor a particularly high frequency of carriers.

Recall that natural selection works to eliminate disadvantageous alleles within the population (purifying selection) and also to promote an increase in frequency of advantageous alleles (positive selection). That occurs because natural selection works though the genetic *fitness* of individuals (their ability to reproduce and have children who survive to a reproductive age): disadvantageous alleles are alleles that reduce fitness; advantageous alleles increase fitness. But sometimes a disadvantageous allele can also simultaneously be an advantageous allele. A form of natural selection called

balancing selection can cause a harmful disease allele to increase in frequency in a population because carriers of the mutant allele have a higher fitness than normal individuals (**heterozygote advantage**).

Sickle-cell anemia provides a classic example of heterozygote advantage. It is very common in populations in which malaria caused by the *Plasmodium falci‑parum* parasite is endemic (or was endemic in the recent past) but is absent from populations in which malaria has not been frequent. In some malaria-infested areas of west Africa, the sickle-cell anemia allele has reached a frequency of 0.15—far too high to be explained by recurrent mutation.

Sickle-cell heterozygotes have red blood cells that are inhospitable to the malarial parasite (which spends part of its life cycle in red blood cells). As a result, they are comparatively resistant to falciparum malaria. Normal homo-zygotes, however, frequently succumb to malaria and are often severely, sometimes fatally, affected. Heterozygotes therefore have a higher fitness than both normal homozygotes and disease homozygotes (who have a fitness close to zero because of their hematological disease).

Heterozygote advantage through comparative resistance to malaria has also been invoked for certain other autosomal recessive disorders that feature hemolytic anemia, such as the thalassemias and glucose-6-phosphate dehydrogenase deficiency. The high incidence of cystic fibrosis in northern European populations and Tay–Sachs disease in Ashkenazi Jews is also likely to have originated from heterozygote advantage, possibly through a greater resistance of carriers to infectious disease.

If continued over many generations, even a small degree of heterozygote advantage can be enough to change allele frequencies significantly (invalidating Hardy–Weinberg predictions that assume constant allele frequencies).

Distinguishing heterozygote advantage from founder effects

Diseases that are common in a population because of a founder effect typically originate from one (or occasionally two) mutant alleles. Most people in the population who carry mutant alleles can be expected to have the same ancestral mutation. Heterozygote advantage, by contrast, could be conferred by multiple different mutations of similar effect in the same gene.

If genealogical evidence is not available, strong support for a founder effect can still be obtained. DNA analyses may show that multiple individuals from the population have alleles with the same pathogenic mutation located within a common haplotype of nonpathogenic alleles at neighboring marker DNA loci. If, by contrast, it can be shown that there are multiple different disease alleles in the population or that the disease alleles are embedded in different haplotypes (suggesting different mutational events), heterozygote advantage is likely to apply. But sometimes it is difficult to distinguish between different possible contributions made by founder effects, heterozygote advantage, and even genetic drift when population sizes are very small.

SUMMARY

- Some human disorders and traits are very largely determined by genetic variation at a single gene locus.
- Multiple members of a human kinship (extended fam ily) may be affected by the same single-gene disorder as a result of genetic transmission of mutant alleles (individual versions of a gene at one locus) from one generation to the next.
- In dominantly inherited disorders, an affected person is usually a heterozygote—one allele at the disease locus is defective or harmful, but the other allele is normal.
- In recessive disorders, an affected person has defec tive alleles only at the disease locus. A person with one disease allele and one normal allele is usually an unaffected carrier who can transmit the harmful (mutant) allele to the next generation.

- A person with an autosomal recessive disorder may have two identical mutant alleles (a true homozygote) or two different mutant alleles (a compound heterozygote).
- Because the X and Y chromosomes have very different genes, men are hemizygous by having a single functional allele for most genes on these chromosomes.
- In X-linked recessive disorders, men are disproportion ately affected (they have a single allele, but women with one mutant allele are usually asymptomatic carriers).
- One X chromosome is randomly inactivated in each cell of the early female embryo; descendant clonal cell populations have an inactivated maternal X or an inactivated paternal X. A female carrier of a mutant X-linked allele may be affected if the normal X has been inactivated in a disproportionately large number of cells.
- A genetic mosaic with a mixture of normal and mutant cells may be mildly affected but transmit the mutant allele to descendants who would have harmful mutations in each cell and be more severely affected.
- Some types of mutation are dynamically unstable and become more severe from one generation to the next (anticipation).
- Affected individuals in the same family can also show differences in phenotype because they have different alleles at some other genetic locus (modifier) that interacts with the disease gene locus.
- Our cells each contain many copies of mitochondrial DNA (mtDNA) and affected individuals in a family with a mtDNA disorder may be variably affected because of heteroplasmy (variable ratios of mutant to normal mtDNA copies per cell).
- In disorders of imprinting, individuals who have inherited a mutant gene may or may not be affected, depending on

whether the mutant allele was inherited from the paternal or maternal line.

- There is no one-to-one correspondence between genes and phenotypes. Different mutations in the same gene can sometimes cause different disorders, and yet the same disorder is quite often caused by mutations in different genes.
- Some single-gene disorders are notably common in certain ethnic populations. For recessive disorders, a high carrier frequency may arise because asymptomatic carriers of the mutant allele have been reproductively more successful than individuals with two normal alleles (the single mutant allele may have given heterozygotes an advantage by providing greater protection against certain infectious diseases).
- Mutant alleles lost from the population (when indi viduals fail to reproduce) are balanced by new mutant alleles (created by fresh mutation). For recessive disorders and late-onset dominant disorders, comparatively few alleles are lost from the population and so fresh mutation rates are low. For a severe dominant disorder that manifests before puberty, the rate of fresh mutation may be very high.
- Allele frequencies can be calculated in populations by using the Hardy-Weinberg law, which gives the frequency of a homozygous genotype as the square of the allele frequency, and the frequency of a heterozygous genotype as twice the product of the allele frequencies.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

Single-gene disorders

McKusick VA (2007) Mendelian inheritance in Man and its online version, OMIM. *Am J Hum Genet* 80:588–604; PMID 17357067.

Pagon RA (eds) *GeneReviews™*. http://www.ncbi.nlm.nih.gov/books/NBK1116/; PMID 20301295 (see also **Box 5.1**).

General Mendelian inheritance

Bennett RL (2008) Standardized human pedigree nomenclature: update and assessment of the recommendations of the National Society of Genetic Counselors. *J Genet Counsel* 17:424–433; PMID 18792771.

Wilkie AOM (1994) The molecular basis of dominance. *J Med Genet* 31:89–98; PMID 8182727.

Zschocke J (2008) Dominant versus recessive: molecular mechanisms in metabolic disease. *J Inherit Metab Dis* 31:599–618; PMID 18932014.

X-linked inheritance and X-inactivation

Franco B & Ballabio A (2006) X-inactivation and human disease: X-linked dominant male-lethal disorders. *Curr Opin Genet Dev* 16:254–259; PMID 16650755.

Mangs AH & Morris BJ (2007) The human pseudoautosomal region (PAR): origin, function and future. *Curr Genomics* 8:129–136; PMID 18660847.

Migeon BR (2007) Females are Mosaics. *X Inactivation and Sex Differences in Disease*. Oxford University Press.

Orstavik KH (2009) X chromosome inactivation in clinical practice. *Hum Genet* 126:363–373; PMID 19396465.

Allele frequencies, mosaicism, and calculating genetic risk

Aidoo M (2002) Protective effects of the sickle cell gene against malaria morbidity and mortality. *Lancet* 359:1311–1312; PMID 11965279.

Clarke A (2019) *Harper's Practical Genetic Counselling*, 8th ed. CRC Press.

Hartl D & Clark AG (2007) *Principles of Population Genetics*, 4th ed. Sinauer Associates.

Hurst LD (2009) Fundamental concepts in genetics: genetics and the understanding of selection. *Nature Rev Genet* 10:83–93; PMID 19119264.

McCabe LL & McCabe ER (1997) Population studies of allele frequencies in single gene disorders: methodological and policy considerations. *Epidemiol Rev* 19:52–60; PMID 9360902.

Van der Meulen MA (1995) Recurrence risks for germinal mosaics revisited. *J Med Genet* 32:102–104; PMID 7760316.

6
Principles of gene regulation and epigenetics

DOI: 10.1201/9781003044406-6

CONTENTS

All our cells develop ultimately from the zygote. Each nucleated cell in a person contains the same set of genes. However, only a subset of the genes in a cell are expressed to make functional end products, and that subset varies according to the type of cell. The global gene expression pattern of a cell dictates the form of a cell, how it behaves, and ultimately its identity—whether it will be a hepatocyte, for example, or a macrophage, or a sperm cell.

In Chapter 2 we outlined the basic details of gene expression. Here, we are concerned with how the expression of genes is regulated. Different levels of gene regulation affect the production or stability of gene products: transcription, post-transcriptional processing (to make final mRNA or noncoding RNA products), translation of mRNA, post-translational modification, folding of protein products, incorporation into a multisubunit functional molecule, and degradation of gene products.

We explore aspects of mRNA degradation and protein folding in Chapter 7. Here we mostly deal with gene regulation at the levels of transcription, post-transcriptional processing, and translation. Complex networks of interacting regulatory nucleotide sequences and proteins are involved.

e two fundamental types of gene regulation

All the cells in our body originate by cell division ultimately from one cell, the fertilized egg cell. Given that in each person the nucleated cells all contain the same DNA molecules, readers might reasonably wonder how we could ever come to have different cell types with distinct gene expression patterns. However, it is not just the sequence of nucleotides in DNA that determines gene expression. Chromatin structure is also crucially important, and gene expression is regulated at two fundamental levels listed below, one of which is not genetic.

- *Genetic regulation*. Here control of gene expression is *dependent* on the nucleotide sequence. If a promoter sequence is deleted, for example, the expected transcript is not produced.
- *Epigenetic regulation*. Here control of gene expression is *independent* of the nucleotide sequence. As detailed in Section 6.2, various non-genetic control mechanisms can affect chromatin structure, causing it to be tightly compacted (preventing expression of genes) or more open (facilitating gene expression).

Major control mechanisms involve certain chemical modifications of DNA and histones, changes in the positioning of nucleosomes, and interactions of certain regulatory noncoding RNAs with chromatin.

We explain in Section 6.2 how epigenetic controls are required in very early development to initiate programs of cell differentiation that progressively leads to different cell lineages and ultimately different cell types. We will also explain how epigenetic controls can be heritable (but can also be reset), and how they can be influenced by environmental and stochastic factors.

Cis-acting and *trans*-acting effects in gene regulation

Genetic control of gene expression largely depends on collections of short regulatory nucleotide sequences in both DNA and RNA; they act as target sequences that can be bound by certain regulatory RNA molecules and proteins.

A regulatory sequence is said to be ***cis-acting*** when its function is limited to the *single* DNA or RNA molecule it resides on. Take gene promoters. The promoter upstream of the insulin gene on a paternally inherited chromosome 11 regulates the *paternal* insulin gene only, not the allelic insulin gene on maternal chromosome 11. In addition, an allele may be regulated by more distantly related *cis*-acting sequences on the same chromosomal DNA molecule (**Figure 6.1A,B**).

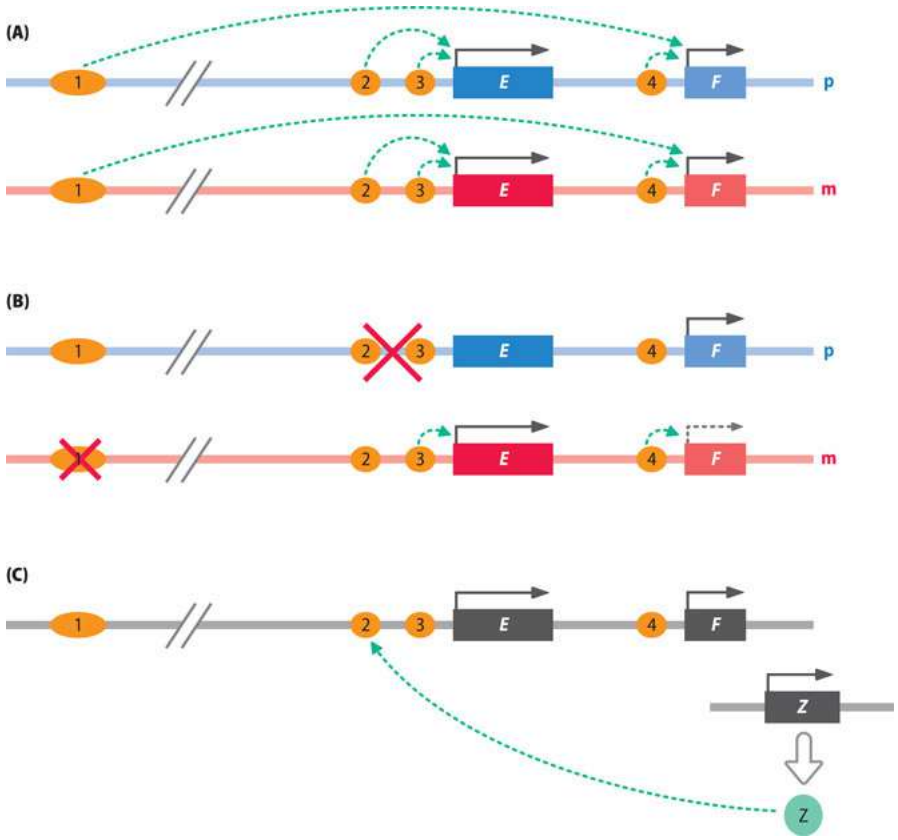


Figure 6.1 Examples of *cis*-acting and *trans*-acting effects at the DNA level. (A) Normal regulation of neighboring genes *E* and *F* on homologous chromosomes by positive *cis*-acting regulatory elements 1 to 4 (orange ovals). Paternal DNA (p) is shown in blue; maternal DNA (m) is in pink. Gene *E* is controlled by elements 2 and 3; gene *F* is controlled by proximal element 4 and remote element 1. (B) Effects of mutation (large red X) abolishing regulatory elements. Deletion of paternal elements 2 and 3 inactivates the paternal *A* allele only; deletion of maternal element 1 selectively reduces expression of the maternal *B* allele. (C) *Trans*-acting gene regulation. A remote gene *Z* on another chromosome (as shown here), or on the same chromosome as genes *E* and *F*, but distantly located (not shown) produces a *trans-acting* regulatory protein Z that binds to regulatory element 2 on *both* paternal and maternal chromosomes (represented here by a single, generic black chromosome).

A ***trans*-acting** gene regulator is a regulatory protein or regulatory RNA molecule that can migrate by diffusion to recognize and bind *specific* short regulatory nucleotide sequences in DNA or RNA. Unlike a *cis*-acting gene regulator, a *trans*-acting gene regulator can regulate the expression of *both* alleles on distantly located genes (Figure 6.1C). And some individual *trans*-acting regulators regulate a set of genes at multiple loci (all of which possess a target nucleotide sequence it can bind to).

Many RNA transcripts also contain *cis-acting* regulatory elements (whose effect is limited to regulating the expression of the RNA transcript on which they reside). Untranslated sequences in mRNA molecules, for example, generally contain *cis*-acting sequences that regulate the expression of the transcript. They may be recognized and bound by *trans*-acting regulatory proteins or *trans*- acting regulatory noncoding RNAs, notably microRNAs (see **Figure 6.2**).

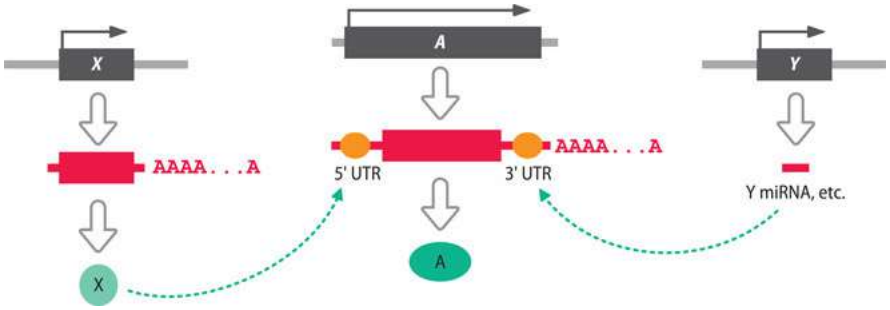


Figure 6.2 ***Cis*-acting and *trans*-acting regulation at the RNA level**. *Cis*-acting regulatory elements in mRNA are often located in untranslated regions (UTRs). Expression of protein-coding gene A is imagined to be regulated at the transcript level by *trans*-acting regulatory protein X and by microRNA Y that bind to *cis*-acting elements in the 5¢ and 3¢ UTRs, respectively, of the mRNA. Note that microRNAs often bind to target nucleotide sequences in RNA transcripts from multiple different genes and thereby regulate the expression of specific sets of genes. Figure 6.9 shows some examples of *transacting* regulatory proteins.

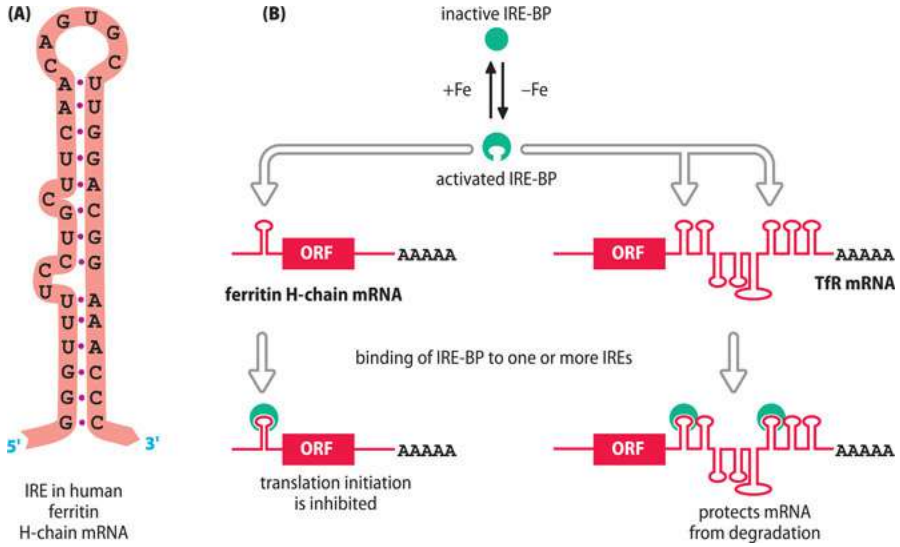


Figure 6.9 Iron-response elements in the ferritin and transferrin mRNAs. (A) Stem-loop structure of an iron-response element (IRE) in the 5¢ untranslated region of the ferritin heavy (H)-chain mRNA. (B) When iron levels are low, a specific IRE-binding protein (IRE-BP) is activated and binds the IRE in the ferritin heavy-chain gene and also IREs in the 3¢ untranslated region of the transferrin receptor (TfR) mRNA. Binding inhibits the translation of ferritin but protects the transferrin receptor mRNA from degradation, maximizing the production of transferrin receptor. When iron levels are high, the IRE-binding protein is inactivated, maximizing the production of ferritin and decreasing the production of transferrin receptor. ORF (open reading frame) designates the central coding DNA of the mRNAs.

Not shown in Figure 6.2 are additional types of *cis*- and *trans*-regulation in which regulatory long noncoding RNA acts on DNA. Unlike *trans*-acting RNA regulators, *cis*-acting RNA regulators are not free to move away by diffusion; instead, they remain within chromatin, attached to the individual DNA strand they were transcribed from. We show in Section 6.2 specific examples of how *cis*-acting regulatory RNA molecules work in epigenetic gene regulation in processes such as X-chromosome inactivation and genome imprinting.

We begin this chapter by looking at how genetic regulation governs how our genes are expressed. We then consider principles of epigenetic regulation. We end with a section on abnormal epigenetic regulation that results from abnormal chromosome inheritance, or from single-gene disorders in which mutations affect a gene involved in epigenetic regulation. We describe a very different type of epigenetic dysregulation that relates to protein folding in Chapter 7, and we examine epigenetic contributions to complex disease in chapters 8 and 10.

6.1 GENETIC REGULATION OF GENE EXPRESSION

As described in Figure 2.12, the mitochondrial genome is transcribed from fixed start points, generating large multigenic transcripts from each DNA strand, which are subsequently cleaved. By contrast, it is usual for nuclear genes to be transcribed individually, and transcription is regulated by genetic factors, as described in this section, and also by epigenetic factors, which we consider in Sections 6.2 and 6.3. Recently, geneticists have also become increasingly aware of the importance of post-transcriptional controls, notably at the level of RNA processing and translation.

Promoters: the major on–off switches in genes

Along the lengths of each DNA strand of our very long chromosomal DNA molecules are **promoters**, *cis*-acting regulatory DNA sequences that are important in establishing which segments of a DNA strand will be transcribed. Each promoter is a collection of very short sequence elements that are usually clustered within a few hundred nucleotides from the transcription start site. For each DNA strand, transcription begins at fixed points on the DNA where the chromatin has been induced to adopt a relaxed, "open" structure (see below).

Nuclear genes are transcribed by three different types of RNA polymerases. A nucleolar RNA polymerase, RNA polymerase I, is dedicated to making three of the four different ribosomal RNAs (rRNAs) in our cytoplasmic ribosomes (the 28S, 18S, and 5.8S rRNAs). It transcribes clusters of about 50 tandem DNA repeats (each containing sequences for the 28S, 18S, and 5.8S rRNAs) on each of the short arms of chromosomes 13, 14, 15, 21, and 22. RNA polymerase II transcribes all protein-coding genes, genes making long noncoding RNAs and some short RNA genes (including many miRNA genes). RNA polymerase III transcribes tRNA genes, the 5S rRNA gene, and some other genes that make short RNAs.

None of the RNA polymerases acts alone; each is assisted by dedicated protein complexes. In the case of RNA polymerase II, for example, a core transcription initiation complex is formed by the sequential assembly of five multisubunit proteins *(general* transcription factors—see below) at specific sites on the DNA.

Some of the protein subunits of the transcription initiation complexes recognize and bind specific short DNA sequence elements of a promoter; others are recruited by binding to previously bound proteins. For a protein-coding gene, most core promoter elements are upstream of the start site, and the spacing of the elements is important.

Figure 6.3 illustrates some important core promoter elements, but note that the composition of core promoter elements is highly variable—some promoters lack all the elements shown in this figure. We describe additional *cis*-acting elements in the next section.

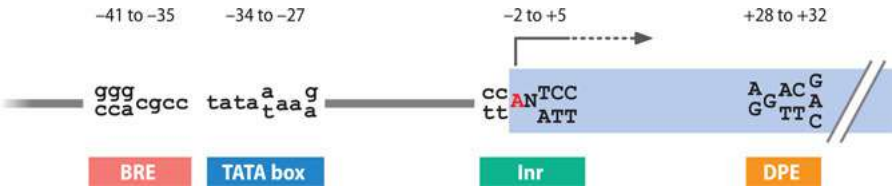


Figure 6.3 Consensus sequences for some core promoter elements often found in genes transcribed by RNA polymerase II. The TATA box is bound by the TATA-binding protein subunit of transcription factor IID. The initiator (Inr) element defines the transcription start site (the highlighted A) when located 25–30 bp from a TATA box. The downstream core promoter (DPE) element is only functional when placed precisely at +28 to +32 bp relative to the highlighted A of an Inr element. TFIIB binds to the BRE (TFII**B r**ecognition **e**lement) and accurately positions RNA polymerase at the transcription start site. However, none of these elements is either necessary or sufficient for promoter activity, and many active polymerase II promoters lack all of them. N represents any nucleotide. (Adapted from Smale ST & Kadonga JT [2003] *Annu Rev Biochem* 72:449–479; PMID 12651739. With permission from Annual Reviews.)

Once the basal transcription apparatus has fully assembled, a component with DNA helicase activity is responsible for locally unwinding the DNA helix, and the activated RNA polymerase accesses the template strand.

Modulating transcription and tissue-specific regulation

As a metaphor for gene expression, imagine the output from a radio. The basal transcription apparatus described above would be the radio's ON switch that is needed to get started. It is required in all cells and uses ubiquitous transcription factors that bind to *cis*-acting elements in the core promoter. But there is also the need for a VOLUME control to amplify or reduce the signal, as required in different cell types or at different stages in a cell's life or development.

The role of the volume control is performed by additional circuits. First, additional, often non-ubiquitous, transcription factors bind to *cis*-acting regulatory elements other than those of the core promoter. These elements are sometimes distantly located from the gene they regulate, as described immediately below. Then there are co-activator or co-repressor proteins that are recruited by bound transcription factors. In addition, diverse types of long noncoding RNAs regulate transcription. Because they often work in the epigenetic regulation of transcription, we consider them in Section 6.2.

Cis-acting regulators as modifiers of basal gene expression

As seen from Figure 6.3 a promoter is made up of sequence elements whose orientation and spacing are important. Two types of additional *cis*-acting regulatory DNA elements modify the transcriptional output in a way that is independent of their orientation:

- **enhancers** amplify transcription;
- **silencers** repress transcription.

Enhancers and silencers may be located close to a transcriptional start site, from shortly upstream of the promoter (of the gene they regulate) to the beginning of its first intron. But quite often, too, they may be rather distantly located from the promoter (**Figure 6.4A**). To allow remote elements such as these to work, the intervening DNA needs to be looped out so that regulatory proteins bound to the enhancer can now physically interact with proteins bound to the promoter of the target gene (see Figure 6.4B).

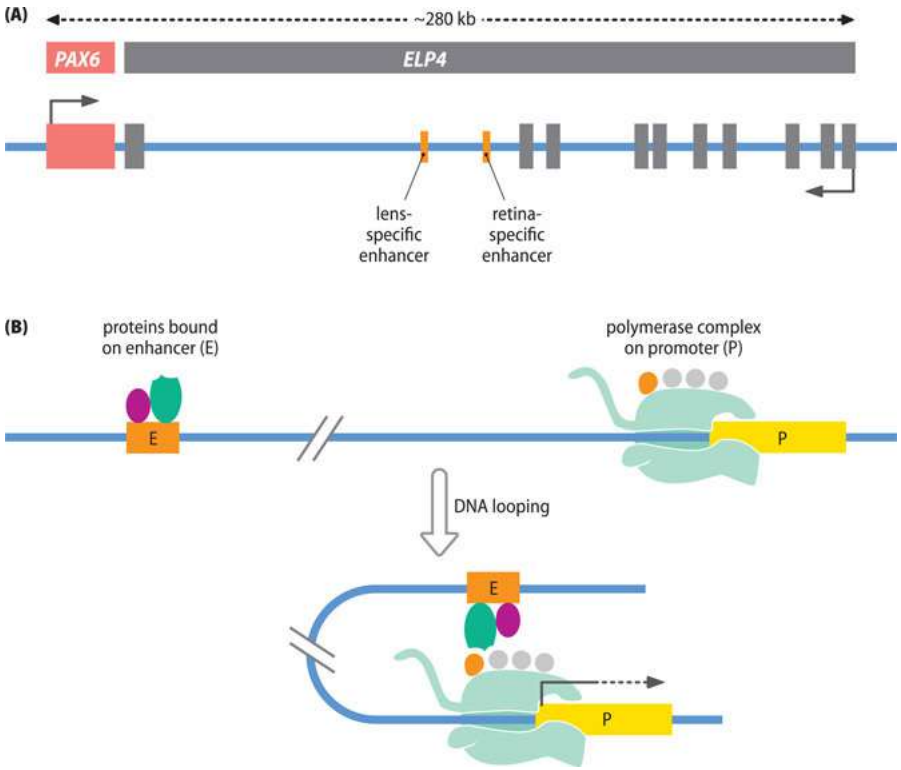


Figure 6.4 DNA looping is required to bring a distant enhancer in close proximity to the promoter of the gene it regulates. (A) Example of remote enhancers. The 33 kb *PAX6* gene (upper left) is mutated in aniridia type I (OMIM 106210) and is known to be regulated by two distantly located enhancers: a lens-specific and a retina-specific enhancer residing within a long intron of the large neighboring *ELP4* gene. Vertical gray boxes represent *ELP4* exons (for clarity, *PAX6* is represented here by a single box representing both exons and introns). Exons and enhancer elements are not to scale. (B) General enhancer-promoter interaction. Bending of the intervening DNA allows direct physical interactions between proteins bound to an enhancer (or other remote *cis*-acting element) with some of the many proteins bound to the promoter of the gene that it regulates. For clarity, only the RNA polymerase is shown at the promoter.

Cis-acting regulators as boundary elements

The long-distance action of elements such as enhancers needs to be targeted to the correct genes. To ensure that signals from regulatory elements do not affect genes other than the intended targets, **boundary elements** are needed to establish physical separation between euchromatin and constitutive heterochromatin, and also between different regions of euchromatin. Two important classes are:

- **barrier elements**. They maintain differences in chromatin structure between neighboring euchromatic and heterochromatic regions.
- **insulators.** They block inappropriate interactions between enhancers and promoters in a region of euchromatin.

We give examples in Section 6.3 to illustrate the use of both types of boundary element.

Transcription factor binding and specificity

A protein **transcription factor** is a sequence-specific DNA-binding protein that binds to specific short target DNA sequences (often four to nine nucleotides long) within or close to genes that it regulates. In addition to ubiquitous general transcription factors (which bind to core promoter elements), many other transcription factors bind to target sequences within additional, often remote, *cis*-acting sequences, such as enhancers, as described below.

Some genes need to be expressed in all cells ("housekeeping" genes), but many are expressed only in specific tissues and/or at specific developmental stages (the activity of the promoters relies on tissue-specific or developmentally regulated transcription factors that bind to noncore elements). Like other DNA-binding proteins, a transcription factor typically recognizes its target sequence using a DNA-binding domain that contains some motifs that physically bind DNA, such as zinc fingers or leucine zippers (Figure 6.5).

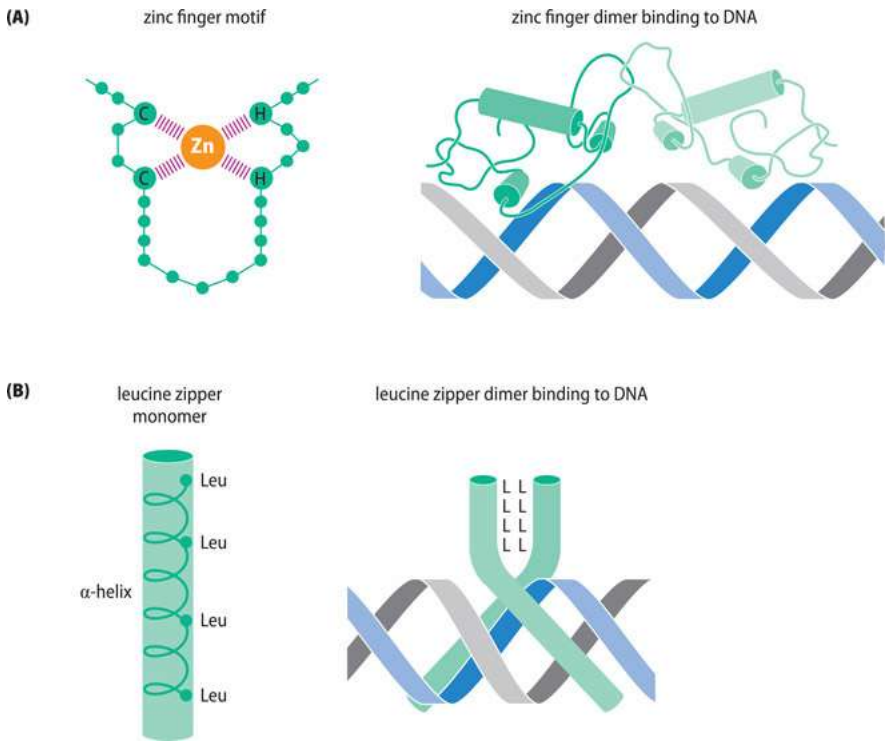


Figure 6.5 Examples of common DNA-binding motifs found in transcription factors and other DNA-binding proteins. (A) In the zinc finger motif a Zn^{2+} ion is bound by four conserved amino acids (normally either histidine or cysteine) to form a loop (finger). Clusters of sequential zinc fingers are common. The so-called C2H2 (Cys_2/His_2) zinc finger typically comprises ~23 amino acids and neighboring fingers are separated by a stretch of about seven or eight amino acids. The structure of a zinc finger may consist of an a-helix and a b-sheet (held together by coordination with the Zn^{2+} ion), or of two a-helices, as shown here. In either case, the primary contact with the DNA is made by an a-helix binding to the major groove. (B) The leucine zipper is a helical stretch of amino acids rich in hydrophobic leucine residues, aligned on one side of the helix. These hydrophobic patches allow two individual a-helical monomers to join together over a short distance to form a coiled coil. Beyond this region, the two a-helices separate, so that the overall dimer is a Y-shaped structure. The dimer is thought to grip the double helix much like a clothes peg grips a clothesline. Leucine zipper proteins normally form homodimers (but can occasionally form heterodimers).

Transcription factor specificity

Transcription factors recognize short target sequences, and so for each transcription factor there are from tens of thousands to hundreds of thousands of potential binding sites across the human genome. Only a tiny fraction of potential target sequences are used, however, for two reasons. First, the binding site must be accessible (in an open chromatin conformation and away from direct contact with nucleosomes). Secondly, transcription factor binding is *combinatorial*—different transcription factors work in concert by binding to adjacent recognition sequences.

A transcriptional activation domain is present in transcription factors that stimulate transcription; transcriptional repressors often recruit specialized protein complexes to silence gene expression, as described in Section 6.2. Other proteins modulate transcription without binding to DNA. Instead, they work by protein–protein interactions

that support other regulatory proteins (which bind DNA directly). There are two types: transcriptional *co-activators* (which enhance transcription) and *co-repressors* (which downregulate transcription).

Genetic regulation during RNA processing: RNA splicing and RNA editing

Understanding the genetic control of splicing is important for understanding pathogenesis because mutations causing abnormal RNA splicing are a relatively common cause of disease. RNA editing is a less well understood form of RNA processing.

Regulation of RNA splicing

Like transcription, RNA splicing is subjected to different controls, and some splicing patterns are ubiquitous; others are tissue-specific. As illustrated in **Figure 6.6A**, three fundamental *cis*-acting regulatory RNA sequences are required for the basic splicing mechanism, which is performed by large ribonucleoprotein complexes known as spliceosomes. The **splice donor site** contains an invariant GU dinucleotide that defines the 5¢ end of an intron at the RNA level. The **splice acceptor site** contains an invariant AG dinucleotide that defines the 3¢ end of an intron at the RNA level and is embedded within a larger sequence that includes a preceding polypyrimidine tract. An additional control element, the branch site, is located very close to the splice acceptor; it contains an invariant A nucleotide and is responsible for initiating the splicing reaction. Note that the sequence surrounding the invariant GU and AG signals is variable—some splice sites are strong and readily used, whereas others are weak and used only occasionally.

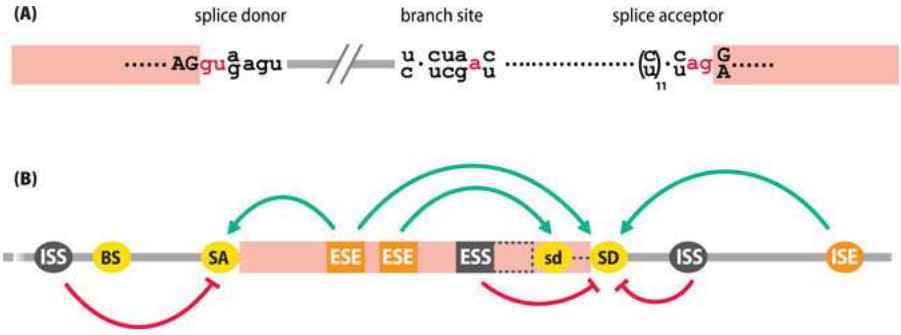


Figure 6.6 ***Cis*-acting sequences that regulate RNA splicing**. Pink boxes represent transcribed exon sequences. (A) The three fundamental RNA target sequences in the splicing mechanism. Bases in red are essentially invariant. The gap shown by the // symbol can vary in length from tens of nucleotides up to several hundred kilobases long in extreme cases. Spliceosomes contain several types of small nuclear RNA (snRNA), including U1 snRNA, which base pairs with the splice donor sequence, and U2 snRNA, which base pairs with the branch site sequence. (B) In addition to the splice donor (SD), splice acceptor (SA), and branch site (BS) sequences, other regulatory RNA sequences stimulate splicing (orange) or inhibit splicing (black). In this example, an exon has two exonic splice enhancers (ESE), and an exonic splice suppressor (ESS) and is flanked by introns that have intronic splice suppressors (ISS). The dotted black lines indicate an alternative 3¢ end to the exon, due to the use of an alternative splice donor (sd) sequence instead of the usual splice donor sequence (SD).

Splicing is also regulated by two additional classes of short (often hexanucleotide) *cis*-acting regulatory RNA elements:

- *splice enhancer* sequences (which stimulate splicing);
- *splice suppressor* sequences (which inhibit splicing).

These sequences are located close to splice junctions and can lie within exons or introns (Figure 6.6B). To help keep the spliceosome in place, splicing enhancers bind SR proteins (so called because they have a domain based on repeats of the serine-arginine dipeptide). Splicing suppressors bind hnRNP proteins that are active in removing

bound spliceosomes. Because different tissues and cell types can express different SR proteins and different hnRNP proteins, splicing patterns can vary between tissues.

Alternative splicing

More than 90 % of human protein-coding genes undergo some kind of alternative splicing, when primary transcripts of a single gene are spliced in different ways (**Figure 6.7** gives some variations, and Figure 6.6B shows how they can be generated). Sometimes, some transcripts retain transcribed intronic sequence. **Exon skipping** occurs when one or more full-length exons are not represented in some transcripts. In other cases, there is some variability in the precise locations of exon–intron junctions, so that transcripts from one gene may have short or long versions of an exon.

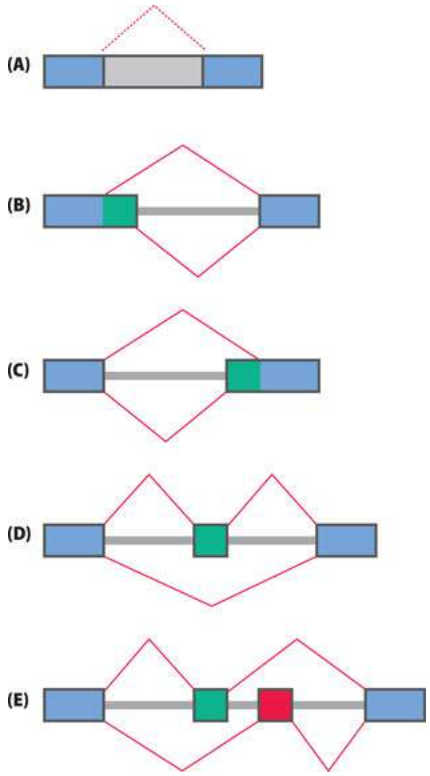


Figure 6.7 Types of alternative splicing event. (A) An intronic sequence (gray) is either excluded from a transcript or retained. (B,C) The use of alternative splice donor sites (B) or of alternative splice acceptor sites (C) results in the inclusion or exclusion of the sequences in green. (D) The exon in green may be either included or skipped (a *cassette exon*). (E) Alternative exons: the mature mRNA includes either the exon in green or the exon in red, but not both or neither. Blue boxes represent exons that are always included in the mature mRNA.

Some of the variable transcripts may be functionally unimportant (splicing accidents must happen occasionally). Often, however, alternative splicing patterns show tissue specificity (so that, for example, one splice pattern is consistently produced in brain but another pattern is normally found in liver), or there may be consistent differences in the use of specific splice patterns at different stages in development. By producing alternative products (**isoforms**) from individual genes, alternative splicing can increase functional variation.

Alternative isoforms may be retained in cells, or secreted, or sent to different cellular compartments (to interact with different molecules and perform different roles). For example, the −KTS isoforms of the WT1 Wilms tumor protein (**Figure 6.8A**) function as DNA-binding transcription factors, but the +KTS isoforms associate with pre-mRNA and may have a general role in RNA splicing. This pattern of alternative splicing has been conserved over hundreds of millions of years. The ERBB4 protein, a tyrosine protein kinase that is a member of the epidermal

growth factor receptor family, has CYT1 and CYT2 isoforms that respectively possess or lack a binding site for the phosphatidylinositol-3-kinase signaling molecule (Figure 6.8A).

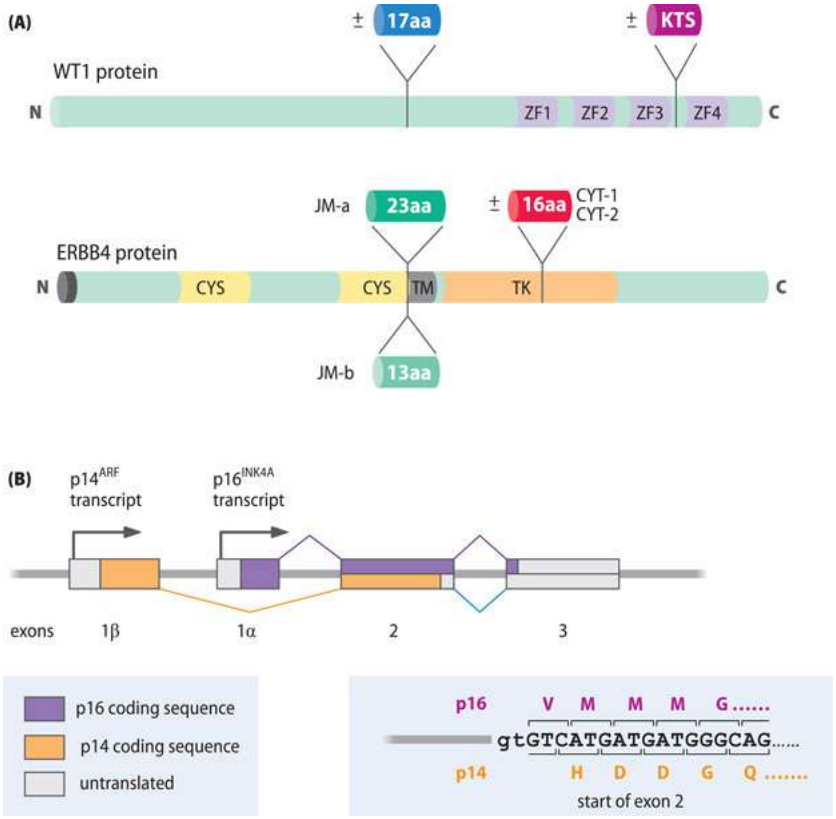


Figure 6.8 Examples of alternative splicing in human genes. (A) Alternative splicing results in the variable presence of a 17 amino acid (17aa) peptide sequence near the middle of the WT1 Wilms tumor protein, and of a Lys-Thr-Ser tripeptide sequence (KTS) between the third and fourth zinc finger (ZF) domains. (B) Four different isoforms also exist for the human ERBB4 protein. Just before the transmembrane (TM) domain there is the alternative presence of a 23aa or a 13aa peptide sequence (JM-a and JM-b isoforms, respectively). Within the tyrosine kinase (TK) domain a 16-amino-acid peptide sequence with a binding site for phosphatidylinositol-3-kinase is present in the CYT-1 isoform but absent in the CYT-2 isoform. (C) Alternative splicing of the *CDKN2A* gene produces two tumor suppressor proteins, p16-INK4A and p14-ARF, that both work in cell cycle control but with entirely different amino acid sequences. Exon 2, the only exon that has a coding sequence for both proteins, is translated in different reading frames (bottom right).

Very occasionally, quite different proteins are created from a common gene by alternative splicing. The *CDKN2A* gene provides a prime example by producing two entirely different proteins that, nevertheless, have similar functions (see Figure 6.8B).

RNA editing

In some RNA transcripts, certain nucleotides naturally undergo deamination or transamination. When this happens in the coding sequences of mRNAs, the amino acid sequence of the protein will differ from that predicted by the genomic DNA sequence. For example, certain adenines in some RNA transcripts are naturally deaminated to give the base inosine (I), which behaves like gua-nine (by base pairing with cytosine). In coding sequences, A ® I editing is most commonly directed at CAG codons, which specify glutamine (Q). The resulting CIG codons behave like CGG and code for arginine (R), and so this type of RNA editing is therefore also called Q/R editing.

Q/R editing is quite commonly found during the maturation of mRNAs that make neurotransmitter receptors or ion channels.

Some other types of RNA editing are known, including C ® U editing (used in making apolipoprotein B mRNA, for example) and U ® C editing (used in making mRNA from the *WT1* Wilms tumor gene). The extent of RNA editing is still controversial, and its significance is unclear.

Translational regulation by *trans*-acting regulatory proteins

Regulation at the level of translation allows cells to respond more rapidly to altered environmental stimuli than is possible by altering transcription. According to need, stores of inactive mRNA may be held in reserve so that they can be translated at the optimal time. Controls are also exerted over where an mRNA is translated: some mRNAs are transported as ribonucleo-protein particles to specific locations within a cell; for example, alternative splicing allows tau mRNA, to be selectively localized to the proximal regions of axons rather than to dendrites.

To control gene expression at the level of mRNA, *trans-acting* regulatory factors bind to specific *cis*-acting RNA elements in the untranslated regions of the mRNA. Single-stranded RNA is quite flexible (unlike DNA, which has a rather rigid structure), but typically it has a very high degree of secondary structure as a result of intra-chain hydrogen bonding (shown in Figures 2.4 and 2.6). RNA elements that bind protein are often structured as hairpins, as in the example of the iron-response element shown in **Figure 6.9A**.

As an example of translational regulation, consider how cells control the availability of two proteins involved in iron metabolism: ferritin (an iron-binding protein used to store iron in cells) and the transferrin receptor (which helps us absorb iron from the diet). When iron levels are low, the priority is to maximize the amount of iron that can be absorbed from the diet: transferrin receptor mRNA is protected from degradation so that it can make a protein product. Conversely, when iron levels are high, ferritin production is activated to store iron in cells. This happens without any change in the production of ferritin or transferrin receptor mRNAs. Instead, both these mRNAs have iron-response elements (IREs), which can be bound by a specific IRE-binding protein that regulates the production of protein from these mRNAs; the availability of this binding protein is also regulated by iron concentrations (see Figure 6.9B).

Post-transcriptional gene silencing by microRNAs

Trans-acting regulators such as the IRE-binding protein that work by binding to mRNA used to be viewed as quirky exceptions. The discovery of tiny RNA regulators, notably microRNA, changed all that. MicroRNAs (miRNAs) are single-stranded regulatory RNAs that downregulate the expression of target genes by base pairing to complementary sequences in their transcripts. Typically about 20–22 nucleotides long, they are formed by multiple processing events, including cleavages in the cytoplasm that are performed by the same endoribonucleases used in RNA interference, a natural cell defense mechanism that we detail in Section 9.4.

A miRNA binds to any transcript that has a suitably long complementary sequence to form a stable heteroduplex (correct base pairing is important for the “seed” sequence covering the first eight or so nucleotides from the 5¢ end of the miRNA; some mismatches are tolerated when the remaining part of the miRNA pairs up). Because miRNAs are short and some base mismatches are tolerated, a single miRNA can regulate transcripts from many different genes (**Figure 6.10**).

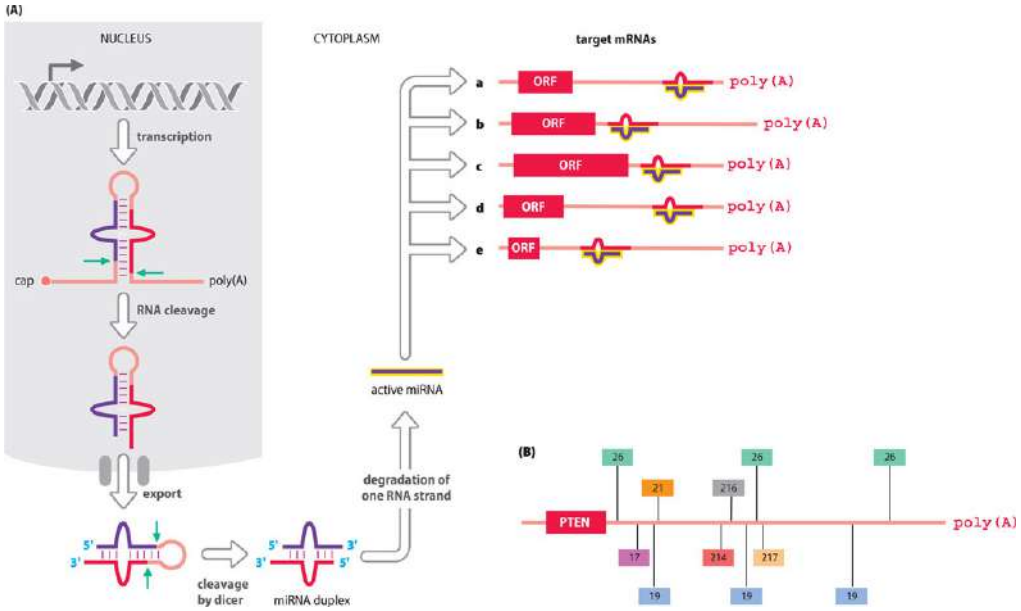


Figure 6.10 How microRNAs are produced and work in cells. (A) miRNA genes are transcribed and cleaved in the nucleus to generate a stem-loop RNA that is exported to the cytoplasm and further cleaved asymmetrically by the endoribonuclease dicer to generate a miRNA duplex with overhanging 3¢ ends. One strand of the duplex (the *passenger strand*, shown in red) is then cleaved and degraded, leaving the other strand (the *guide strand*, shown in purple) as a mature single-stranded miRNA. A typical human miRNA binds to and regulates transcripts produced by hundreds of different genes, and the vast majority of miRNA–target RNA heteroduplexes have imperfect base pairing. Shown here for illustration are five mRNAs produced from five different genes (a–e); the complementary sequence to which the miRNA binds is shown in red. ORF, open reading frame (= coding DNA). (B) An individual mRNA often has multiple miRNA-binding sites. The example here is the mRNA from the human *PTEN* tumor suppressor gene that has binding sites in the 3¢ untranslated region for miRNAs belonging to seven miRNA families: three binding sites each for miR-19 and miR-26, and one each for miR-17, miR-21, miR-214, miR-216, and miR-217.

We have several hundred miRNA genes, and they frequently show tissue-specific expression; many are important in early development, but miRNAs have been found to be important regulators in a whole range of different cellular and tissue functions. At least 50 % of our protein-coded genes are thought to be regulated by miRNAs, and individual types of mRNA often have recognition sequences for multiple miRNA regulators. Just like protein transcription factors, miRNAs seem to be involved in complex regulatory networks, and they are subject to negative regulation by a wide range of RNA classes as described in the next section.

Repressing the repressors: competing endogenous RNAs sequester miRNA

Many of our pseudogenes are known to be transcribed. Some of them seem to have undergone purifying selection, indicating that they are functionally important. A landmark study published in 2010 provided the first real insights into how functional pseudogenes work: it showed that the human *PTEN* gene at 10q23 is regulated by a highly related processed pseudogene, *PTENP1*, located at 9p21.

PTEN makes a protein tyrosine phosphatase that is very tightly controlled (cells are very sensitive to even subtle decreases in abundance of this protein, and aberrant *PTEN* expression is common in cancers). *PTENP1* does not make a protein (one of the changes from the *PTEN* sequence disrupts the initiator methionine codon). It does, however, make a noncoding RNA that retains many of the miRNA-binding sites in the 3¢ UTR of *PTEN* mRNA. The *PTENP1* RNA seems to regulate *PTEN* expression by binding to and sequestering miRNAs that would normally bind to the *PTEN* mRNA (Figure 6.11 gives the principle).

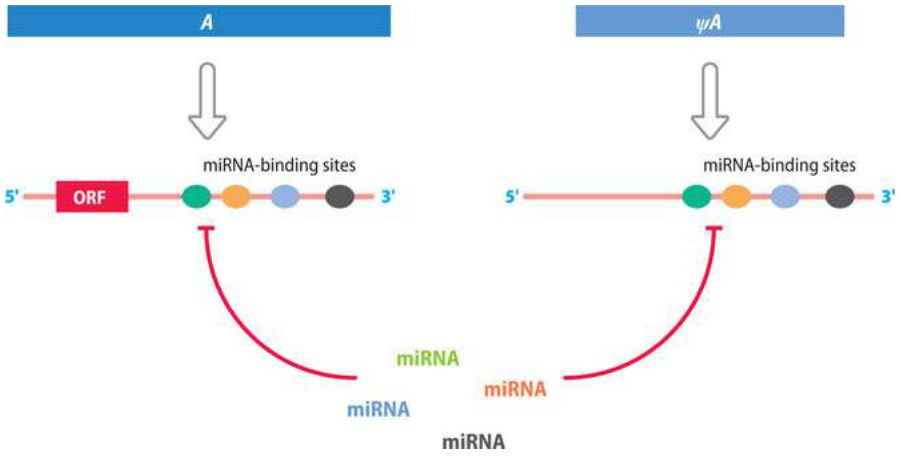


Figure 6.11 **Different types of competing endogenous RNAs can act as miRNA sponges**. In this example, a protein-coding gene A and a closely related pseudogene *ψA* produce RNA transcripts that have in common binding sites for certain miRNA classes. The pseudogene RNA can compete with the mRNA for binding by the same miRNA classes (by soaking up miRNA the *ψA* RNA acts as a "miRNA sponge"). Various other RNA classes can also act as miRNA sponges, including certain long noncoding RNAs and also *circular RNAs*. The latter are surprisingly abundant in our cells and form by head-to-tail splicing reactions ("back-splicing") that join first and last exon sequences; since they largely overlap protein-coding sequences they can contain sequences corresponding to the untranslated sequences of mRNAs, including miRNA-binding sites.

6.2 CHROMATIN MODIFICATION AND EPIGENETIC FACTORS IN GENE REGULATION

During development, cell differentiation (and the production of different cell lineages) is dictated by programmes of altered gene expression that are independent of the DNA sequence. Instead, they depend on altered epigenetic settings (often called **epigenetic marks**) that affect chromatin structure (and thereby gene expression). In its broadest sense, **epigenetics** covers all phenomena that can produce heritable changes in how genomes function without affecting the base pairing properties of the DNA sequence.

An overview of the molecular basis of epigenetic mechanisms

Later in this section we provide detail on some individual classes of epigenetic mechanisms. First, we briefly outline the characteristics of five important classes. Three comparatively well understood classes involve certain types of chemical modification of DNA or of histones bound to DNA, plus substitution of standard histones by histone sequence variants. Nucleosome positioning and *cis*-acting regulatory non-coding RNAs are also important (see Table 6.1).

TABLE 6.1 FIVE IMPORTANT CLASSES OF EPIGENETIC MECHANISMS AFFECTING CHROMATIN STRUCTURE

Epigenetic mechanism	Comments
DNA methylation	Specifically, methylation of cytosines within CpG dinucleotides to give 5-methylcytosine (which base pairs like cytosine). The palindromic nature of CpG provides a simple way of transferring this epigenetic mark to daughter DNA strands (Figure 6.15). Chromatin with highly methylated DNA is condensed, and transcriptionally inactive but hypomethylation is associated with an open chromatin structure (Figure 6.12).

* The term *chromosome remodeling* encompasses repositioning of nucleosomes and histone substitution.

Epigenetic mechanism	Comments
Histone modification	Post-translation chemical modification of side chains occurs at multiple relatively accessible amino acids on the C-terminus tails of histones. Three common types of modification are: acetylation (at certain lysines); methylation (at certain arginines and lysines); and phosphorylation (mostly directed at certain serines and threonines)—(Table 6.3).
Histone substitution*	The replacement of standard histones by certain other histone sequence variants (Table 6.4).
Nucleosome repositioning*	Chromatin modeling complexes are large ATP-powered multiproteins that physically drive nucleosomes along the DNA to create areas of low or high nucleosome density.They help set up active or repressed chromatin states (an active transcription site typically has ~ 150 bp of nucleosome-free DNA, and highly ordered flanking nucleosomes).
Cis-regulation by noncoding RNA	These regulatory RNAs remain attached to the DNA they were transcribed from. They act either as antisense RNAs or as scaffolds for binding regulatory protein complexes to change chromatin structure. See below for important examples in imprinting and X-inactivation.

* The term *chromosome remodeling* encompasses repositioning of nucleosomes and histone substitution.

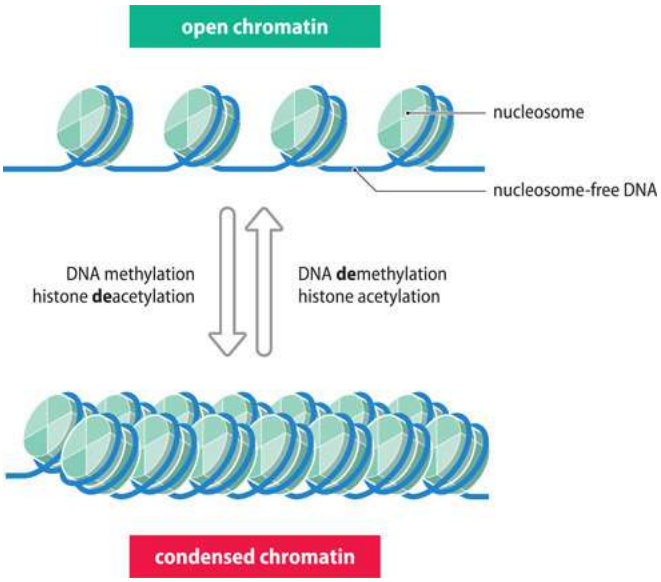


Figure 6.12 Altered chromatin states arise from DNA and chromatin modifications. In chromatin that has an open configuration the nucleosome-free stretches of DNA may include promoter elements and other regulatory sequences to which regulatory factors can bind, allowing gene expression. But in highly condensed (compacted) chromatin the transcription factors are denied access by the tight packing. Note that whereas DNA methylation (and histone deacetylation) is associated with condensed chromatin, different types of histone methylation are associated with open and condensed chromatin (see Table 6.3).

TABLE 6.3 EXAMPLES OF HISTONE MODIFICATIONS CHARACTERISTIC OF DIFFERENT CHROMATIN STATES*

	EUCHROMATIN					HETEROCHROM	
	Promoters		Enhancers		Gene bodies	Facultative	Const
AMINOACID	Active	Inactive	Active	Inactive	Inactive		
H3K4	H3K4me2/me3		H3K4me1/me2				

* The nomenclature for histone modifications gives first the histone class, then the amino acid in one-letter code, followed by the position of the amino aci (counting from the N-terminus), and finally the chemical modification. So, for example, H3K9Ac represents acetylation of the lysine at the 9th amino acid counting from the N-terminus of histone H3.

	EUCHROMATIN					HETEROCHRON	
	Promoters		Enhancers		Gene bodies	Facultative	Const
AMINOACID	Active	Inactive	Active	Inactive	Inactive		
H3K9	H3K9ac	H3K9me3		H3K9me2/me3	H3K9me2/me3	H3K9me2	H3K9
H3K27	H3K27ac	H3K27me3	H3K27ac			H3K27me3	
H4K12	H4K12ac					H4K12ac	H4K1
H4K20							H4K2

* The nomenclature for histone modifications gives first the histone class, then the amino acid in one-letter code, followed by the position of the amino aci (counting from the N-terminus), and finally the chemical modification. So, for example, H3K9Ac represents acetylation of the lysine at the 9th amino acid counting from the N-terminus of histone H3.

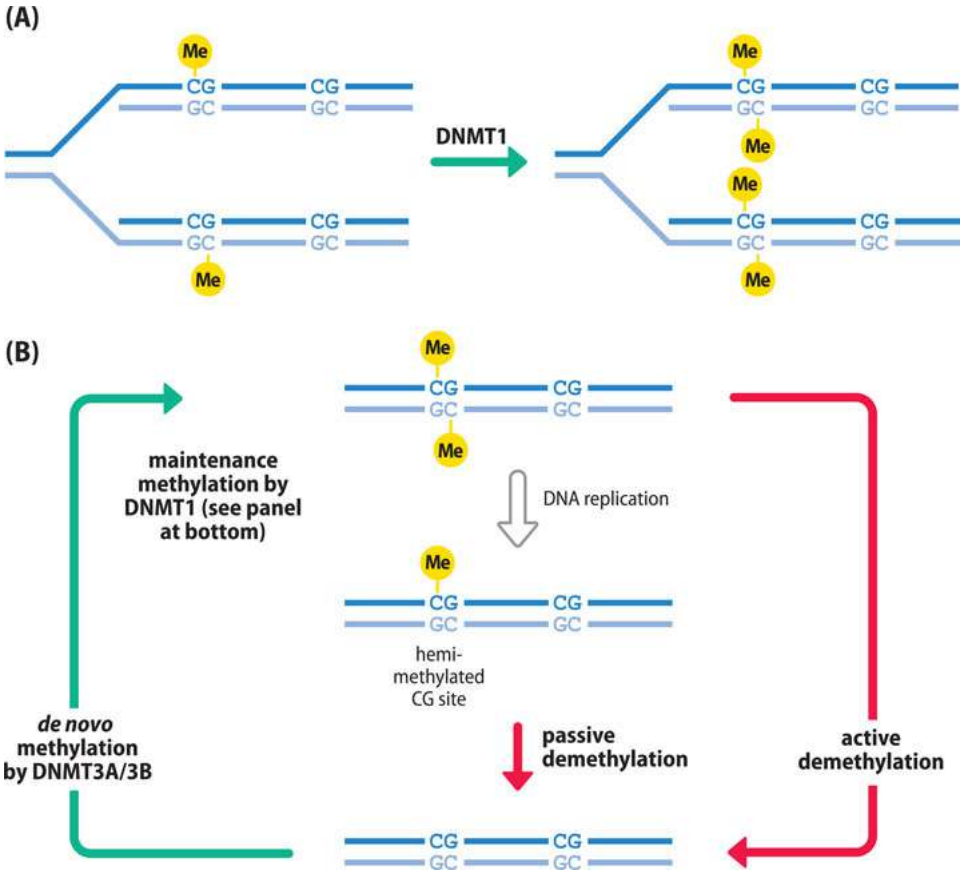


Figure 6.15 DNA methylation and demethylation mechanisms in mammalian cells. (A) Maintenance methylation. During replication of a DNA molecule containing methylated CG dinucleotides, the parental strand retains methylated cytosines, but the newly synthesized DNA incorporates unmodified cytosines. DNMT1 is usually available, and it specifically methylates any CG dinucleotides on the newly synthesized strand that are paired with a methylated CG on the parental strand, regenerating the original methylation pattern. (B) Pathways towards DNA methylation (green arrows) and demethylation (red arrows). If DNMT1 is not available, the hemimethylated DNA can give rise in a subsequent DNA replication to unmethylated DNA (passive demethylation). Unmethylated DNA can also be generated by an active demethylation process at certain stages in development. DNMT3A and DNMT3B are used for *de novo* methylation at specific developmental stages (see Figure 6.16).

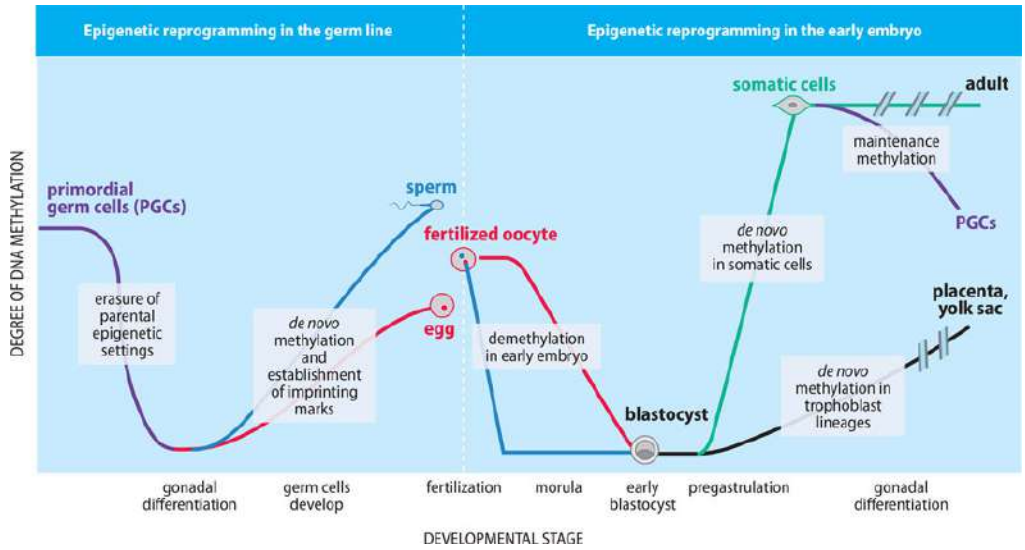


Figure 6.16 Changes in DNA methylation during mammalian development. Marked and often tissue-specific changes in overall methylation accompany gametogenesis and early embryonic development. Their causal role remains uncertain, although mice that are specifically unable to methylate sperm DNA are infertile. The horizontal time axis is necessarily abbreviated on the right-hand side of the figure leading toward birth and then adulthood (indicated by the use of double slashes) PGCs, primordial germ cells.

TABLE 6.4 EXAMPLES OF HISTONE H2A AND HISTONE H3 VARIANTS

Class	Variant	Description
H2A	H2AX	important in DNA repair and recombination (it is introduced at sites of double-strand breaks)
	H2A.Z	associated with the promoters of active genes. It also helps prevent the spread of silent heterochromatin and is important in maintaining genome stability
H3	H3.3	important in transcriptional activation
	CENP-A	a centromere-specific variant of H3. It is required for assembly of the kinetochore, to which spindle fibers attach

Heritability of epigenetic marks

Epigenetic marks can be stably transmitted from one cell generation to the next, providing a form of cellular memory. For example, once a cell has been epigenetically programmed to become an intestinal epithelial cell, daughter cells retain this programming so that they, too, are intestinal epithelial cells. And in addition to regulating how genes are expressed, epigenetic settings determine how some DNA sequences determine chromosome functions. DNA sequences at centromeres and telomeres, for example, have heritable epigenetic settings, and these sequences will continue to dictate centromere and telomere functions in the daughter cells. The patterns of constitutive heterochromatin in a parent cell are also reproduced in daughter cells.

There are some instances in nature, notably in plants, in which epigenetic effects can be transmitted through meiosis, from one organism to subsequent generations. In mammals, however, major waves of epigenetic reprogramming occur in gametogenesis and in the early embryo to remove parental DNA methylation marks and reset the global DNA methylation patterns (see **Table 6.2**). As a result, epigenetic marks are not normally transmitted across generations from parents to children (but we describe later some limited evidence in Chapter 8).

TABLE 6.2 EXAMPLES OF EPIGENETIC PHENOMENA INVOLVING DNA AND CHROMATIN MODIFICATION IN MAMMALIAN CELLS

Phenomenon	Mechanism/comments

Phenomenon	Mechanism/comments
Epigenetic reprogramming in gametogenesis	Readily detected as a wave of genome wide demethylation during germ cell development (erasing parental epigenetic marks) followed by comprehensive *de novo* DNA methylation to reset global patterns of DNA methylation and gene expression.
Epigenetic reprogramming in the early embryo	Eggs and sperm are differentiated cells, and their genomes have different epigenetic marks. They combine to give a zygote whose genome is gradually reprogrammed to erase the great majority of the inherited epigenetic marks. By the blastocyst stage, the cells of the inner cell mass are pluri potent and will give rise ultimately to all cells of the body. Epigenetic marks are reestablished in the descendants of the cells of the inner cell mass to establish different cell lineages and permit cell differentiation.
Establishment of centromeric heterochromatin	Centromere establishment relies on nucleosomes incorporating a specific histone H3 variant known as CENP-A.
X-chromosome inactivation	Initiated by the *XIST* long noncoding RNA that somehow coats most of one of the two X chromosomes in female cells, silencing most of its genes.
Genomic imprinting	Silencing of one allele, according to parent of origin, at diverse gene loci (often organized in gene clusters) on different chromosomes.
Position effects causing heteroch romatin ization	Large-scale changes in DNA, causing genes to be relocated to a region of constitutive heterochromatin where they are silenced.

Stability of epigenetic marks

Although epigenetic marks are often stable, they can be changed. A naturally occurring example is evident during germ cell development and in the very early embryo where epigenetic marks in mammalian genomes are programmed to be reset across the genome between generations. (And, of course, they can be reset artificially to clone animals, as in the case of the famous sheep Dolly, and in cultured cells to create induced pluripotent stem cells, for example.)

Some epigenetic marks can also be reset naturally in response to environmental conditions. Cells receive a wide range of extracellular chemical signals, notably from neighboring cells but also from chemicals in food that we ingest.

How changes in chromatin structure produce altered gene expression

Binding of DNA to a histone octamer and bending of the DNA on the surface of the histone octamer to form regular nucleosomes make it very difficult for regulatory factors such as transcription factors to bind to their target sequences. Depending on its chromatin environment, the properties of a DNA sequence can change. A functional gene when embedded in highly condensed chromatin may not be accessible to transcription factors and would be *silenced*. But if the chromatin structure is altered, adopting a more open, relaxed conformation, protein factors may be able to bind the promoter and related regulatory sequences to initiate transcription.

Sometimes a normally expressed gene is transposed (by a translocation or inversion) so that it takes up a new position within, or close to, a region of constitutive heterochromatin (permanently condensed heterochromatin). When that happens the gene would be silenced (an example of a ***position effect***). For mammalian cells, the most striking evidence that gene expression is dependent not only on DNA sequence but also on chromatin structure comes from the X chromosome. In females one X chromosome is very highly condensed across nearly all its

length and genes are silenced across most of the chromosome, unlike in the other X chromosome which has a comparatively open structure.

DNA methylation and chemical modification of histones are important regulators of chromatin structure. DNA methylation involves adding methyl groups to a small percentage of the cytosines and demethylation of DNA involves removing methyl groups from some of the methylated cytosines. (Note: because methylated cytosines behave like cytosine and base pair with guanines, the base sequence is not considered to be altered.) Extensive methylation of DNA sequences in vertebrates is generally characteristic of tightly packed chromatin; loosely structured chromatin ("open" chromatin) has low-level DNA methylation (**Figure 6.12**).

Histone modifications include different types of post-translational modification at specific amino acid positions on the different types of histone. Histone acetylation, for example, is associated with open chromatin, and histone deacetylation with condensed chromatin (see Figure 6.12).

The need for chromatin writers, erasers, and readers

Many different enzymes are responsible for creating or interpreting epigenetic marks, and belong to three classes as follows:

- "*writers*" add chemical groups to modify DNA or histones covalently; in the latter case, different enzymes are employed according to the chemical group deposited, and also according to the numerical position of the amino acid in the histone tail
- "*erasers*" work in the opposite direction to remove the chemical groups
- "*readers*" are involved in binding to specific chemical groups on DNA or histones to interpret defined epigenetic marks.

The readers may recruit additional factors to induce different changes in chromatin, such as chromosome compaction, or changes in nucleosome spacing and structure (**chromatin remodeling**). By adjusting the position of nucleosomes with respect to the DNA strand, promoters and other regulatory DNA sequences can become nucleosome-free, allowing access by transcription factors.

Histone modification and histone substitution in nucleosomes

Nucleosomes have 146 bp of DNA wrapped around a core of eight histone proteins, composed of two each of four different histone classes: H2A, H2B, H3, and H4. The histone proteins are positively charged (having multiple lysine and argi-nine residues) and have protruding N-terminal tails. Although the histone tails in **Figure 6.13A** are shown in isolation, they can make contact with adjacent nucleosomes.

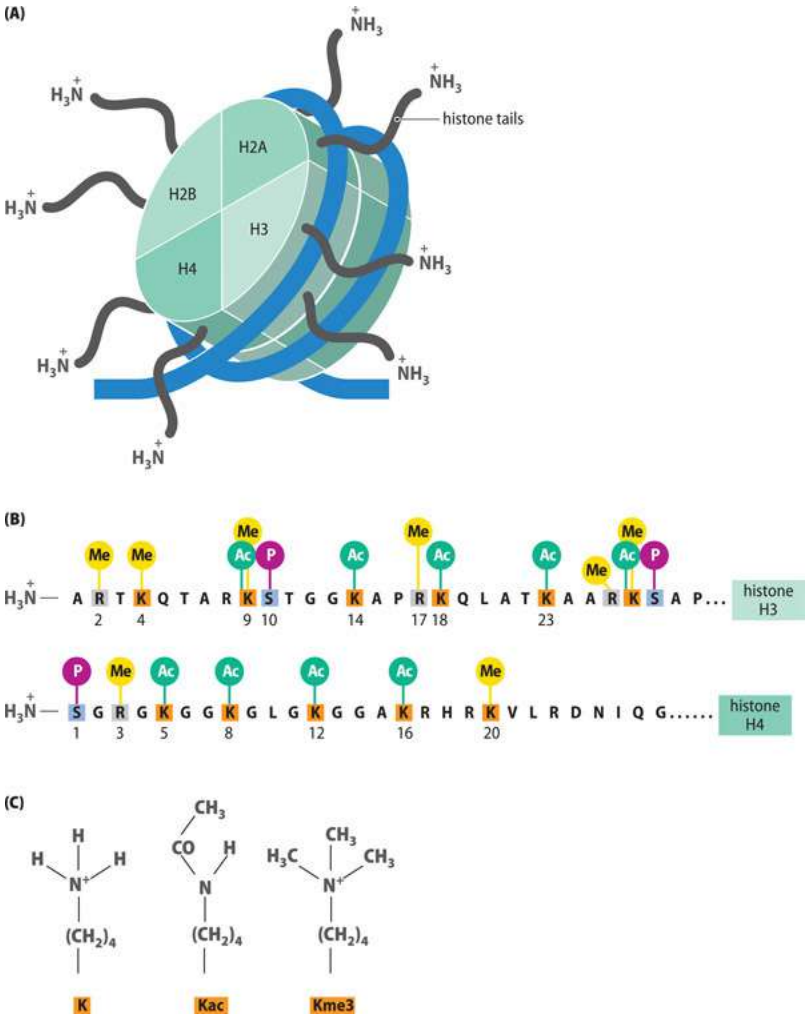


Figure 6.13 Nucleosome structure and histone modifications. (A) Positively charged N-terminal histone tails protrude from nucleosomes and can associate with other nucleosomes (not shown). (B) Map of histone H3 and H4 tail modifications. Note that lysine at position 9 in H3 can be methylated (H3K9me) or acetylated (H3K9ac), but never both. Some lysine residues may have two or three methyl groups (see Table 6.3). (C) Examples of lysine modification, showing the standard side chain of lysine (K), an acetylated lysine (Kac), and a trimethylated lysine (Kme3).

Each N-terminal histone tail shows a pattern of variable chemical modifications at specific amino acid positions. Individual amino acids in each tail may be methylated, acetylated, or phosphorylated (see Figure 6.13B), or subject to yet other types of modification, including tagging with ubiquitin. Particular types of amino acid are preferred targets for modifying the N-terminal histone tails: acetylation occurs only at lysine residues, phosphorylation mostly occurs at serines, but both lysine and arginine residues can be methylated.

Acetylation of lysines leads to loss of the positive charge (Figure 6.13C), and as a result acetylated histone tails interact less well with neighboring nucleosomes than do the unacetylated histone tails. Histone acetylation therefore results in a more relaxed chromatin conformation. According to the specific amino acid position, lysines may also be modified to contain one, two, or three methyl groups (but in these cases the positive charge is retained on the side chain—see Figure 6.13C).

Histone modifications are performed by a series of different enzymes that are devoted to adding or removing a chemical group at specific amino acid positions. Thus, for example, there are multiple histone acetyl transferases (HATs) and histone deacetylases (HDACs), and suites of histone lysine methyltransferases (KMTs) and histone lysine demethylases (KDMs).

Histone substitution

Nucleosome structure is also altered by histone substitution: standard his-tones in nucleosomes are substituted by minor *histone variants* that recruit regulatory factors. As described below, this can have different effects, such as activating transcription or defining centromeres. Although much of our knowledge of chromatin modification has come from studying patterns of DNA methylation and histone modification or substitution, numerous nonhistone proteins and noncoding RNAs also have important roles in modifying chromatin structure.

Modified histones and histone variants affect chromatin structure

Histone modifications are "read" by nonhistone proteins that recognize and bind the modified amino acids and then recruit other proteins to effect a change in chromatin structure. Proteins with a bromodomain recognize the acetylated lysines of nucleosomal histones, those with a chromodomain recognize methylated lysines, and different varieties of each domain can recognize specific lysine residues. Chromatin-binding proteins often have several domains that recognize histone modifications.

Certain individual types of histone modification are associated with open chromatin and transcriptional activation, or with condensed chromatin and transcriptional repression. For example, methylation of H3K4 (the lysine at position 4 on histone H3) is associated with open chromatin at the promoters of actively transcribed genes and at active enhancers (**Table 6.3**). By contrast, trimethylation of the lysine at position 9 on histone H3 (H3K9me3) is prominently associated with transcriptional repression, being widely found in constitutive heterochromatin and in inactive genes in euchromatin (see Table 6.3).

In addition to histone modification, core histone proteins can be replaced by minor variants, notably of histone classes 2A and 3 (the variants typically differ from the canonical histone by just a few amino acids). The minor histone variants are synthesized throughout interphase and are often inserted into previously formed chromatin by a histone exchange reaction powered by a chromatin remodeling complex. Once inserted, they recruit specific binding proteins to effect some change in the chromatin status for specific functions. A well-studied example is CENP-A, a centromere-specific histone H3 variant that is responsible for assembling kinetochores at centromeres. **Table 6.4** gives other examples.

Modified histones and histone variants typically work together with DNA methylation and demethylation in regulating gene expression (**Figure 6.14**). H3K9me3 can bind heterochromatin protein 1, which in turn recruits DNA methyltransferases that also serve to repress transcription. In turn, DNA methyltransferases and 5-meCG-binding proteins recruit histone deacetylases and appropriate histone methyltransferases to reinforce transcriptional repression.

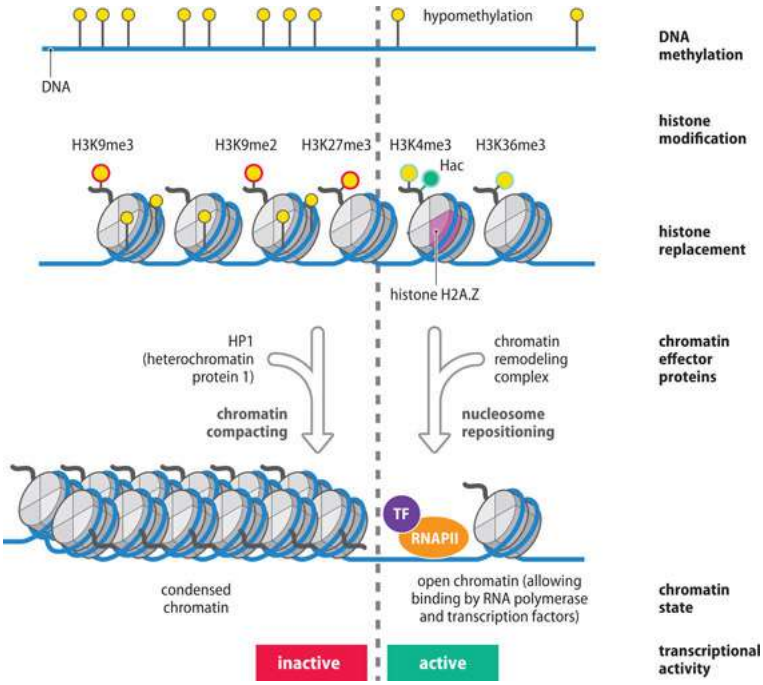


Figure 6.14 Contributions made by DNA and histone modification to different chromatin states in neighboring regions on a chromosome. For convenience, each nucleosome is shown with only one out of the eight N-terminal histone tails. Methylation modifications are shown by circles filled in yellow, comprising small circles (DNA methylation) and large circles (histone methylation). Note that some histone methylations (red outer lines) repress transcription, but other types (green outer lines) are associated with transcription. The filled green circle labeled Hac denotes histone acetylation (which applies to multiple lysines on H3 and H4). Chromatin effector proteins such as heterochromatin protein 1 (HP1)—or the repressive Polycomb group protein complexes PRC2 and PRC1 (not shown here)—are recruited to bind to specific histone modifications, often through the involvement of long noncoding RNAs. Chromatin remodeling can involve creating nucleosome-free regions of DNA that allows the binding of transcription factors. RNAPII, RNA polymerase II.

The function of DNA methylation in mammalian cells

Histones undergo very many different types of chemical modification, as do the nucleotide bases of RNA. However, only one major type of chemical modification is normally seen in DNA: methylation of certain cytosines, a chemical modification that is essential for mammalian development.

The principal function of DNA methylation is to regulate gene expression—it stabilizes, or locks in, patterns of gene silencing so that transcription is suppressed in highly methylated regions of chromatin. Highly repetitive DNA sequences, such as satellite repeats in pericentromeric heterochromatin and dispersed transposons, are extensively methylated. As detailed below, there is also significant—though more sporadic—methylation in the main body of genes (exons and introns) and in intergenic regions.

In the case of DNA methylation patterns, it is the extent of DNA methylation in the key *cis*-acting regulatory elements that distinguishes actively transcribing genes from silenced genes. Thus, the promoters and enhancers of actively transcribing genes are relatively free of DNA methylation. Along with characteristic histone modifications and histone variants such as H2A.Z and H3.3, such hypomethylated regions signal a locally open chromatin environment—transcription factors can gain access to and bind to their target sequences to stimulate transcription. Gene silencing is achieved when significant levels of DNA methylation in a gene's promoter and enhancer elements cause the chromatin to be condensed, denying access to transcription factors.

Like other epigenetic marks, global DNA methylation patterns vary between different cell types and different stages in development. In addition to a general role in gene expression silencing, DNA methylation has important

roles in genomic imprinting and C-chromosome inactivation—we consider these epi-genetic phenomena in more detail below.

DNA methylation is also important for suppressing retrotransposon elements. Retrotransposons are evolutionarily advantageous to genomes because they can insert varied DNA sequences at different locations in the genome, potentially providing novel exon combinations in genes (see Figure 2.16) and giving birth to new regulatory sequences and new exons. About 43 % of the human genome is made up of retrotransposon repeats, but the number of actively transposing retrotransposons needs to be carefully regulated to prevent the genome from being overwhelmed. By suppressing retrotransposon transcription, DNA methylation acts as a necessary brake on excessive transposon proliferation.

DNA methylation: mechanisms, heritability, and global roles during early development and gametogenesis

In mammalian cells, DNA methylation involves adding a methyl group to certain cytosine residues, forming 5-methylcytosine (5-meC). The cytosines that are methylated occur within the context of a palindromic sequence, the CG dinucleotide (also called a CpG dinucleotide, where p represents phosphate).

The 5-meC base pairs normally with guanine (the methyl group is located on the outside of the DNA double helix, minimizing any effect on base pairing). It is recognized by specific 5-meCG-binding proteins that regulate chromatin structure and gene expression, as described below. In a somatic cell, about 70–80 % of CG dinucleotides will have a methylated cytosine, but the pattern of methylation is variable across the genome and across genes (**Box 6.1**).

BOX 6.1 CpG ISLANDS AND PATTERNS OF DNA METHYLATION ACROSS THE GENOME AND ACROSS GENES

DNA methylation is generally used to “lock in” transcriptional inactivity in regions of our cells that do not require expression. Accordingly, heterochromatin and intergenic regions are subject to high levels of DNA methylation. Hypermethylation of some regions, such as pericentromeric heterochromatin, is important for genome stability; a significant decrease in methylation levels in these regions can lead to mitotic recombination and genome instability. Our genes are also subject to DNA methylation; however, by comparison with heterochromatin and intergenic regions, the DNA methylation is generally reduced and more variable.

As described in the text, DNA methylation in our cells is limited to cytosines and occurs within the context of the dinucleotide CG (because CG is the target for cytosine methylation). The resulting 5-methylcytosine can undergo spontaneous deamination to give thymi-dine (Figure 4.3), and during vertebrate evolution there has been a steady erosion of CG dinucleotides. As in other vertebrate genomes, therefore, the dinucleotide CG is notably under-represented in our DNA (41 % of our genome is made up of G–C base pairs, giving individual base frequencies of 20.5 % each for G and C; the expected frequency of the CG dinucleotide is therefore 20.5 % × 20.5 % = 4.2 %, but the observed CG frequency is significantly less than 1 %).

Within the sea of our CG-deficient DNA are nearly 30 000 small islands of DNA in which the CG frequency is the expected value but cytosine methylation is suppressed. Such **CpG islands** (or CG islands; the p signifies the phosphate connecting C to G) are often 1 kb or less in length and are notably associated with genes. Approximately 50 % of CpG islands are located in the vicinity of known transcriptional start sites, as illustrated in **Figure 1**. A further 25 % of CpG islands are found in the main gene body.

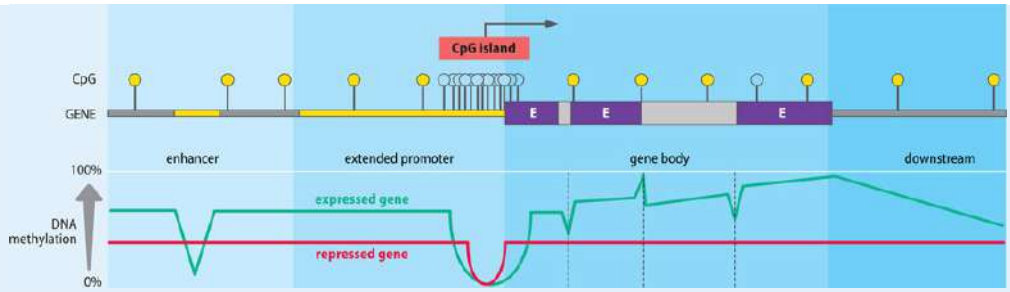


Figure 1 CpG density and DNA methylation levels across an idealized human gene. In this example we consider a gene that has single CpG island located in the vicinity of the transcriptional start site (marked by an arrow) and three exons (E). The CpG density across the gene is illustrated at the top, with open circles indicating unmethylated CpG dinucleotides and circles filled in yellow representing methylated CpG. If the gene is expressed, the gene body shows quite high DNA methylation levels, but the transcriptional start site and upstream enhancer are free of cytosine methylation, allowing access for *trans*-acting protein factors. Even when the gene is transcriptionally inactive, the cytosines remain unmethylated within the CpG island. (Adapted from Hassler MR & Egger G [2012] *Biochimie* 94:2219–2230. With permission from Elsevier Masson SAS.)

Note that CpG islands associated with transcriptional start sites remain unmethylated even when the gene is not being transcribed. Whether a gene is silenced or expressed seems instead to be related to the methylation status of other CpGs that are often located up to 2 kb from transcriptional start sites in CpG island "shores".

DNA methylation mechanism

DNA methylation is performed by DNA methyltransferases (DNMTs). The DNMT1 enzyme serves to maintain an existing DNA methylation pattern. During replication of a methylated DNA molecule, each parental DNA strand retains its pattern of methylated cytosines. The newly synthesized complementary DNA strand is formed by incorporating unmethylated bases, and so in the absence of any further DNA methylation the result would be a hemimethylated DNA (**Figure 6.15A**).

To maintain the original DNA methylation pattern, DNMT1 is normally present at the replication fork and methylates the newly synthesized DNA strand. It methylates only those CG dinucleotides that are paired with a methylated CG on the opposing parental DNA strand. That is, the methylated parent DNA strands act as templates for copying the original methylation pattern (see Figure 6.15A). As a result, patterns of symmetric CG methylation can be faithfully transmitted from parent cell to daughter cells.

The enzymes DNMT3A and DNMT3B are *de novo* methyltransferases—they can methylate any suitable CG dinucleotide (see Figure 6.15B). They have important roles in epigenetic reprogramming when epigenetic marks are comprehensively reset across the genome, at two major stages. In each case, a wave of global DNA demethylation is followed by *de novo* DNA methylation that establishes a different methylation pattern.

DNA methylation in early development and gametogenesis

Major epigenetic reprogramming occurs in the early embryo. Egg cells and, notably, sperm cells have extensively methylated DNA (but rather different patterns of DNA methylation). Following fertilization, the introduced sperm genome (now within the male pronucleus) begins to undergo active DNA demethylation; after the male and female pronuclei fuse, global demethylation of the zygote begins and continues until the early blastula stage in the preimplantation embryo (**Figure 6.16**). Then a wave of genome re-methylation occurs, coincident with initial differentiation steps giving rise to different cell lineages. Genome methylation is extensive in somatic cell lineages but moderate in trophoblast-derived lineages (which will give rise to placenta, yolk sac, and so on).

Significant epigenetic reprogramming also occurs during gametogenesis. The **primordial germ cells** that will give rise ultimately to gametes are initially heavily methylated. As they enter the genital ridge, their genomes are progressively demethylated, erasing the vast majority of epigenetic settings (see Figure 6.16). Thereafter, *de novo* methylation allows epigenetic marks to be reset.

Long noncoding RNAs in mammalian epigenetic regulation

Diverse noncoding RNAs (ncRNAs) have important roles in gene regulation. Mammalian miRNAs are focused on post-transcriptional gene silencing (often by binding to specific target sequences in the untranslated regions of messenger RNAs). Although endogenous siRNAs are important in the epigenetic regulation of centromeres in some organisms, it is not clear that they have a similar role in mammals (if they do, it might be limited to gametogenesis or early embryo development).

The most recent GENCODE release (version 40, April 2022) gives a total of 18 805 human long noncoding RNA genes that have been identified, almost as many as the 19 988 human protein-coding genes. Although many of the long noncoding RNAs have not been well studied, a large proportion are retained in the nucleus and are associated with chromatin. Many of these genes are believed to play roles in chromosome architecture and gene regulation.

We describe later how specific long noncoding RNAs have critical roles in epi-genetic phenomena such as X-chromosome inactivation and imprinting. Here we describe some basic details about how long noncoding regulatory RNAs work.

Cis-acting and *trans*-acting long noncoding RNAs

Unlike microRNAs which regulate genes at the post-transcriptional level, many noncoding RNAs work within chromatin, and have roles in gene regulation. Two major classes are listed below.

- *Antisense RNAs*. These *cis*-acting RNAs can interfere with transcription of partially or fully overlapping genes on the opposite DNA strand, thereby downregulating them. The natural antisense RNAs may be unspliced or spliced, and may silence several genes in a cluster—see **Table 6.5** for examples.

TABLE 6.5 EXAMPLES OF DIFFERENT CLASSES OF REGULATORY LONG NONCODING RNAS

Mode of action	Example	Characteristics
trans-acting gene repression	HOTAIR	2.2 kb RNA encoded from within the *HOXC* cluster at 12q13 and represses *HOXD* genes at 2q31
cis-acting gene repression through antisense RNA (which remains tethered to the DNA strand from which it was transcribed)	KCNQ10T1 (LIT1)	92 kb unspliced antisense RNA that represses transcription of the *CDKN1Cgene* on the opposite DNA strand at 11p15
	SNHG14 (SNRPN)	460 kb spliced and polyadenylated antisense RNA; part of the sequence is antisense to the *UBE3A* gene and represses it.
	CDKN2B-AS1 (ANRIL)	3.8 kb spliced and polyadenylated antisense RNA that represses transcription of *CDKN2B* gene at 9p31; also recruits PRC2* to silence co-located genes

* PRC2, Polycomb repressive protein complex 2 (initiates gene silencing and then recruits the PRC1 complex to maintain it).

Mode of action	Example	Characteristics
cis-acting gene activation by recruiting a chromatin-activating protein complex	HOTTIP	3.8 kb RNA that works by inducing DNA looping to bring target genes into close proximity and then recruiting the WDR5-MLL1 protein complex to deposit transcription-activating H3K4me epigenetic marks
cis-acting gene repression by recruiting a chromatin-repressing protein complex	XIST	19 kb RNA that initiates X-chromosome inactivation; represses transcription of most genes on the inactive X by recruiting the PRC2* complex to deposit transcription-repressing H3K27me3 epigenetic marks

* PRC2, Polycomb repressive protein complex 2 (initiates gene silencing and then recruits the PRC1 complex to maintain it).

- *Chromatin-modifying long noncoding RNAs*. After binding to their target genes (the genes they regulate), many long noncoding regulatory RNAs recruit chromatin-modifying protein complexes to the vicinity, allowing them to change the chromatin status of the target genes. Often long non-coding RNA bound to its target gene attracts a repressive protein complex to the chromatin to bring about local chromosome compaction to silence the target gene; some, however, work to activate transcription of a target gene—see Table 6.5.

As an example, the Polycomb repressive complex 2 (PRC2) is often recruited by regulatory long noncoding RNAs to repress their target genes. It has a methyltransferase subunit that deposits the H3K27me3 epigenetic mark associated with facultative heterochromatin. Note that PRC2 cannot bind to chromatin directly; instead, it needs to be recruited to chromatin by a regulatory RNA that is able to bind both to chromatin and also to PRC2. After the repressive chromatin state has been initiated by PRC2, it is maintained by another Polycomb repressive complex 1 (PRC1).

Figure 6.17 outlines the process for PRC2-induced repression of chromatin for *trans*-acting and *cis*-acting long noncoding RNAs. In the latter case, the long noncoding RNAs may work as newly synthesized RNA transcripts that are still physically associated with the chromatin where they act as scaffolds for assembling repressive protein complexes.

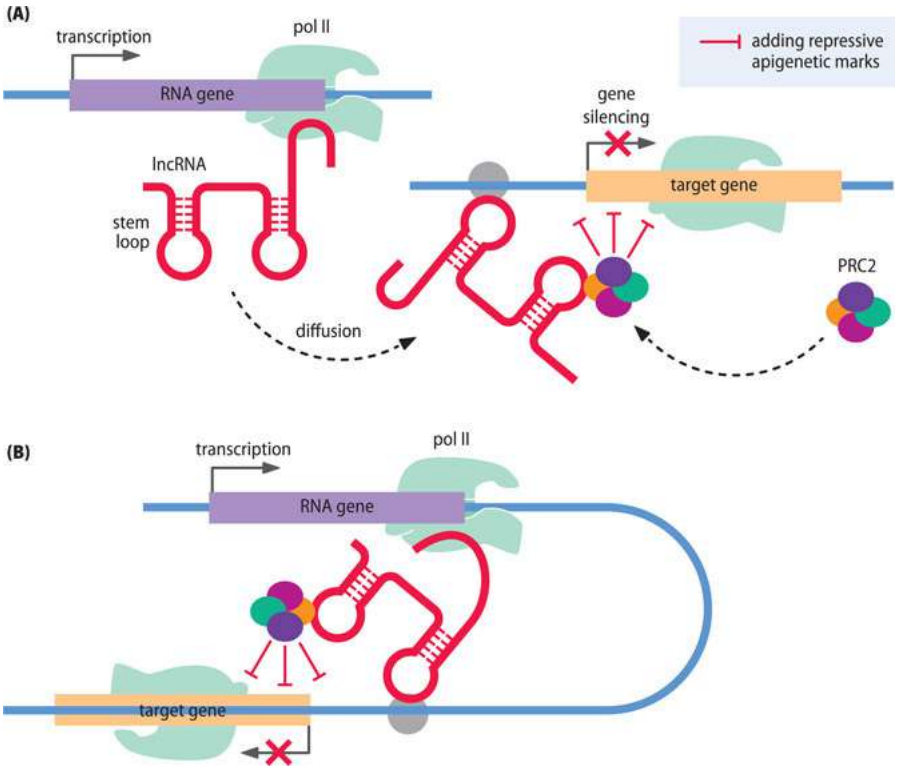


Figure 6.17 How *trans*-acting long noncoding RNA (A) and *cis*-acting long noncoding RNA (B) can silence target genes by recruiting a repressive protein complex such as the PRC2 Polycomb repressive complex. The long noncoding RNA (lncRNA) is envisaged to have two important binding sites (represented here, illustratively, as stem loops). One of these is for binding to the target gene (or to a sequence-specific binding protein bound to the gene, as shown here), and the other is for binding the PRC2 complex. The PRC2 complex cannot bind to target genes by itself. Instead, it relies on a long noncoding RNA with a PRC2-binding site that will bind it, and thereby position it next to a target gene. (A) For *trans*-acting lncRNAs, the RNA is synthesized, migrates to find and bind a target gene (often on a different chromosome), and then recruits PRC2 to deposit its epigenetic mark to silence the gene. (B) For *cis*-acting lncRNAs the newly synthesized RNA remains attached to the chromatin but can arrange (by DNA looping in this case) to bind to a nearby target gene on the same chromosome and recruit PCR2 to silence it.

Genomic imprinting: differential expression of maternally and paternally inherited alleles

We are accustomed to the idea that in mammalian diploid cells both the paternal and maternal alleles are expressed (biallelic expression). For a signifi-cant proportion of our genes, however, only one of the two alleles is normally expressed—the other allele is silenced (monoallelic expression).

Monoallelic expression at a locus can occur at random in an individual: in some cells the paternal allele of a gene is silenced, and in other cells the maternal allele is silenced (Table 6.6). However, for some genes a paternal allele is consistently silenced or a maternal allele is consistently silenced. That is, silencing of one allele occurs according to the parent of origin (**genomic imprinting**).

TABLE 6.6 MONOALLELIC EXPRESSION IN MAMMALS CAN OCCUR BY DIFFERENT MECHANISMS

Class	Mechanisms	Comments
Dependent on parent of origin	genomic imprinting	several genes are expressed only from paternally inherited chromosomes and several only from maternally inherited chromosomes

* At least in eutherian mammals (in marsupials the paternal X is consistently inactivated). Ig, immunoglobulin.

Class	Mechanisms	Comments
	X-inactivation in placenta	paternal X is always inactivated
Independent of parent of origin	X-inactivation in somatic cells*	inactivation of most genes on an X chromosome chosen at random, either the paternal or maternal X (see Figure 6.20A)
	production of *cell-specific* Ig and T-cell receptors	each mature B and T cell makes, respectively, Ig or T-cell receptor chains, using only one allele at a time. Once a functional chain is made by gene rearrangement at one randomly selected allele, a feedback mechanism inhibits further rearrangements (see Section 4.4)
	production of *cell-specific* olfactory receptors	each olfactory neuron expresses a single allele of just a single olfactory receptor (OR) gene (selected from several hundred OR genes) so that it fires in response to one specific odorant only. Depends on competition for a single monoallelic enhancer
	stochastic mechanisms	may be quite common

* At least in eutherian mammals (in marsupials the paternal X is consistently inactivated). Ig, immunoglobulin.

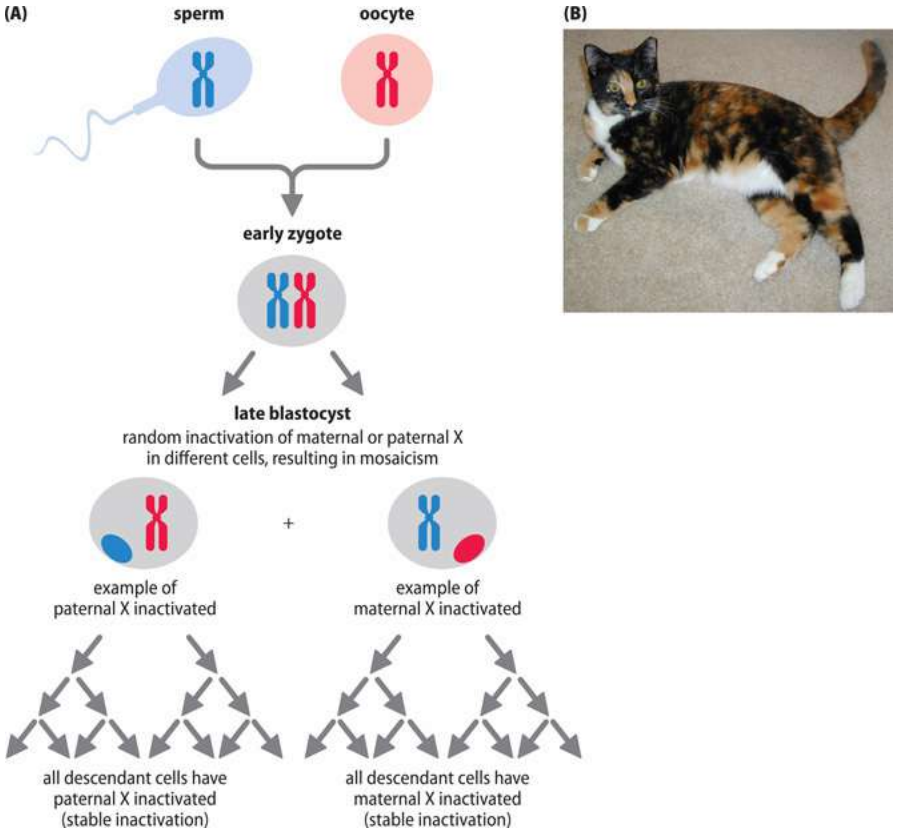


Figure 6.20 X-chromosome inactivation. (A) A randomly chosen X chromosome, either the maternal X or the paternal X, is inactivated in each cell of a 46,XX embryo. Once the choice is made, it is faithfully transmitted through all subsequent rounds of mitosis. Note that in the oogonia of a female, both X chromosomes are active; each has an equal chance of being passed on through the egg. (B) The coat of a calico cat has a mixture of white patches and two other colors, often orange and black as in this case. Like a tortoiseshell cat, it is heterozygous at an X-linked coat color locus, and in the cat represented here one allele specifies a black coat color, the other orange. The different color patches reflect clones in which different X chromosomes are inactivated. The white patches are the result of an unrelated coat color gene. (Adapted from Migeon BR [1994] *Trends Genet* 10:230–235; PMID 8091502. With permission from Elsevier.)

Because natural monoallelic expression occurs for a significant fraction of genes, the maternal and paternal genomes are not functionally equivalent in mammals. As a result, unlike several vertebrate species, mammals cannot naturally reproduce by *parthenogenesis* (reproduction without fertilization: a haploid chromosome set simply duplicates within the oocyte). It is possible to manipulate mammalian eggs artificially to make a diploid embryo with two maternal genomes, but embryonic lethality always ensues: the maternal genome cannot by itself support development—both a maternal and a paternal genome are required. Parthenogenesis fails in mammals because a subset of developmentally important genes is expressed only if inherited paternally; a different subset of genes is expressed only if inherited maternally.

As detailed below, imprinting patterns in genes are established by *cis*-acting regulatory sequences located at an *imprinting control region* that carries a type of imprint, often being methylated to very different extents in sperm DNA and egg DNA. The same differentially methylated region (DMR) can behave very differently when hypomethylated or extensively methylated.

A single allele can behave differently according to the parent of origin, but within an individual the pattern of transcriptional activity or inactivity is maintained through mitosis when somatic cells divide. The alleles are not intrinsically maternal or paternal, however: imprints need to be reversible. A man can inherit an allele from his mother that is inactive. But when he transmits that same allele in his sperm to the next generation the imprint needs to be erased. The allele is reactivated. which can result in reversal of imprints between generations (see **Figure 6.18**).

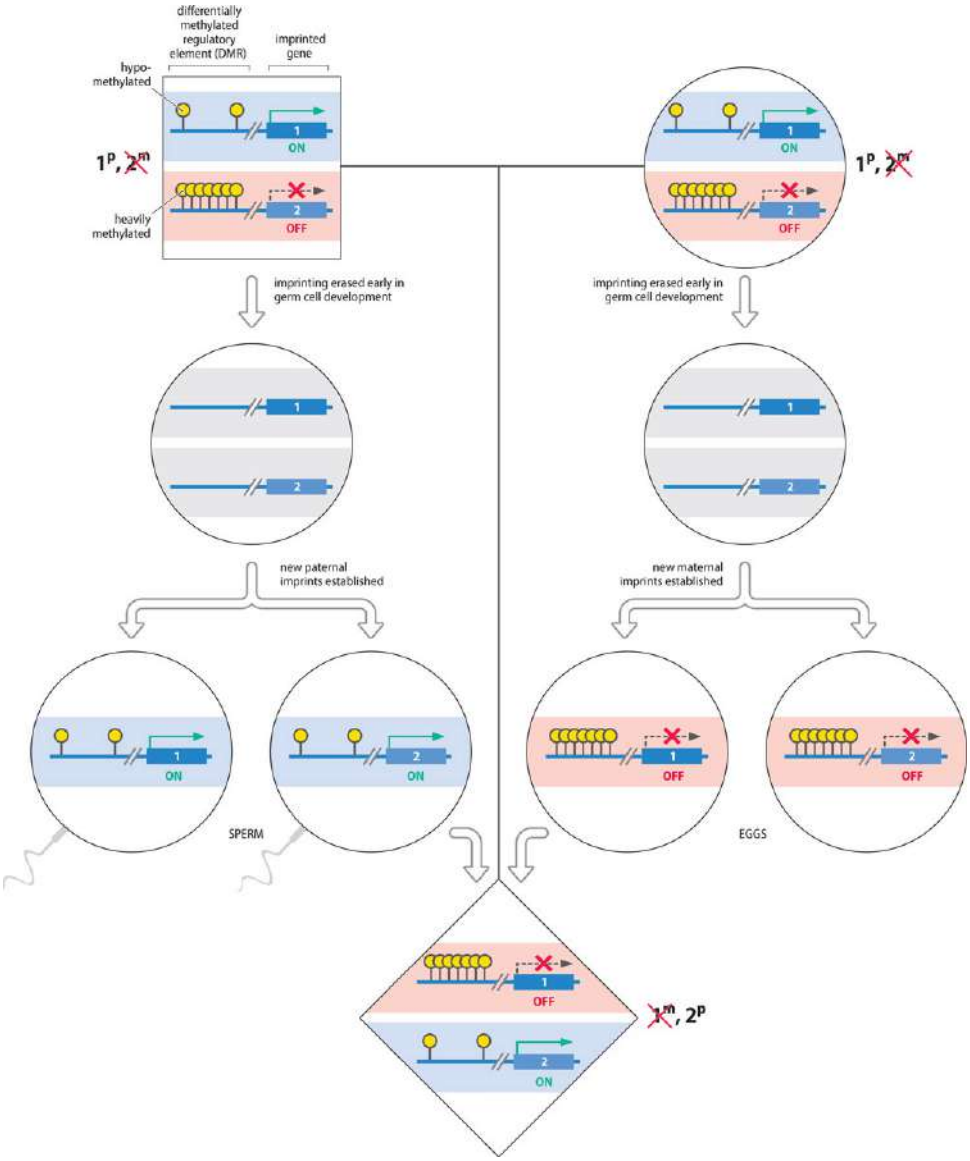


Figure 6.18 How imprints can be reversed between generations. In this example, a differentially methylated regulatory element (DMR) is heavily methylated when inherited maternally, so that a neighboring imprinted gene is silenced (OFF; red cross indicates transcriptional inactivation). When inherited paternally, the DMR is hypomethylated, and the gene is expressed (ON). The gene has two alleles, 1 and 2, and the man and woman at the top are both heterozygotes. Each of them has inherited an active allele 1 on a paternal chromosome (1^{p}) which is illustrated in pale blue shading; and an inactive allele 2 on a maternal chromosome (2^{m}) which is illustrated by pink shading. During early germ cell development, rapid demethylation of DNA occurs, and paternal and maternal epigenetic settings are erased (see left part of Figure 6.16). The loss of parental imprints, including DNA methylation marks, means that the chromosomes mostly lose their parental epigenetic marks and are now represented in neutral gray shading (middle of figure). Later in germ cell development, new imprints are established according to the sex of the individual. Each sperm from the man has an active allele 1 or an active allele 2; each of the woman's eggs has an inactive allele 1 or an inactive allele 2. Fertilization of an egg bearing allele 1 by a sperm carrying allele 2 can generate a child with the same genotype (1,2) as the parents, but the imprint in this child has been reversed: allele 1 is now the inactive allele (having been maternally inherited) and allele 2 is functional.

Extent and significance of genome imprinting

More than 140 mouse genes have been experimentally shown to be imprinted genes, and a smaller number of human imprinted genes have been validated. Catalogs of imprinted genes in different species are available at the geneimprint database at https://www.geneimprint.com/site/genes-by-species.

Many known imprinted genes have a role in embryonic and placental growth and development, and a popular theory attributes imprinting to a conflict of evolutionary interest between mothers and fathers. Propagation of paternal genes would be favored if the offspring were all very robust, even at the expense of the mother (potentially, a man can father children by very many different mothers). Enhanced propagation of maternal genes, however, depends on the mother's being healthy enough to have multiple pregnancies.

Mammalian development is unusual in that the zygote gives rise to both an embryo and also extra-embryonic membranes (including the trophoblast; these membranes act to support development, giving rise to the placenta). From the arguments above, paternal genes might be expected to promote the growth (and general robustness) of the fetus by maximizing the nutrients it can extract from the mother via the placenta. Paternal genes might therefore have a vested interest in supporting the development of the extra-embryonic membranes and placenta. Maternal genes, by contrast, might seek to limit the nutrient transfer so that it does not compromise the mother's health and future reproductive success. Some support for the paternal-maternal conflict theory comes from rare cases of uniparental diploidy in humans (**Figure 6.19**) and from artificially induced uniparental diploidy in mice.

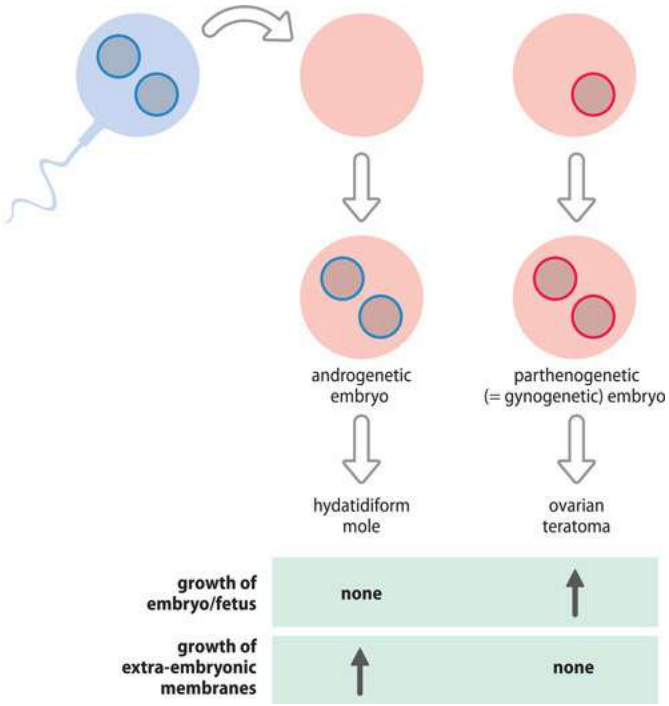


Figure 6.19 Uniparental diploidy and the divergent expression of paternal and maternal genomes. On rare occasions, a zygote is formed with a genome composed of paternal DNA only, producing an *androgenetic embryo* (this usually occurs when a diploid sperm fertilizes a faulty egg that lacks chromosomes, as shown here). Development produces an abnormal conceptus known as a hydatidiform mole, with widespread hyperplasia (overgrowth) of the trophoblast but no fetal parts. The reverse situation, where the zygotic genome is composed of maternal DNA only—producing a *pathenogenetic embryo*—gives rise to an ovarian teratoma that consists of disorganized embryonic tissues without the vital extra-embryonic membranes.

The paternal-maternal conflict theory can only be a partial explanation of why imprinting evolved in mammals. Not all imprinted genes have roles in intrauterine growth, and not all are imprinted in the direction predicted by the parental conflict theory. The theory also doesn't explain why imprinting is tissue-specific for many genes. The insulin-like growth factor gene *IGF2,* for example, is maternally imprinted in many tissues but biallelically

expressed in brain, adult liver, and so on, and the *UBE3A* gene, which is implicated in Angelman syndrome, is paternally imprinted in neurons but biallelically expressed in glial cells and other tissues.

Establishing sex-specific imprints by differential methylation

Imprinted genes are often found in clusters and under the control of a common differential methylation region known as an imprinting control region (ICR). During germ cell development, parental imprints are erased. Thereafter, imprinting is established when sex-specific patterns of *de novo* DNA methylation are created within CG dinucleotides in the ICRs.

Although the sperm and egg have very different overall DNA methylation patterns, large-scale demethylation in the early embryo removes the vast majority of the DNA methylation differences in the paternal and maternal genomes. An important exception is in the ICRs, where the sex differences in DNA methylation are retained in somatic cells (but will be erased in primordial germ cells).

Much of our knowledge of how ICRs regulate the expression of imprinted genes comes from studying certain imprinted gene clusters that are associated with developmental disorders in humans. We consider these mechanisms in the context of disorders of imprinting in Section 6.3; as we will see, long noncoding RNAs are also often associated with clusters of imprinted genes and are important in imprinting control.

X-chromosome inactivation: compensating for sex differences in gene dosage

As described in Section 7.5, a change in constitutional chromosome number (aneuploidy) is usually lethal, or it results in significant abnormalities. Chromosome loss is particularly damaging, and monosomy is lethal in the early embryo with just one exception: loss of an X chromosome from a female embryo is viable and results in Turner syndrome (45,X).

Aneuploidies cause problems because of abnormal **gene dosage**. We have elaborate gene interaction systems, and the products of genes across the genome sometimes can work together in ways where the relative amounts of participating gene products need to be very tightly controlled (changes in the amounts of an individual component can wreak havoc with regulatory systems, for example). An average chromosome has multiple genes in which a change in copy number (producing abnormal gene dosage) is positively harmful. And yet there is one glaring difference that is somehow tolerated in mammals: females have two X chromosomes, but males have only one X chromosome and a Y chromosome.

Whereas Y-specific genes are rare and largely devoted to male-specific functions, the X chromosome has more than 800 protein-coding genes and many RNA genes that work in all kinds of important cell functions. As first proposed by Mary Lyon, a gene dosage compensation mechanism equalizes X-chromosome gene expression in male and female cells by causing one of the two X chromosomes in female cells to be heterochromatinized (**X-chromosome inactivation**).

X-chromosome counting and inactivation choices

Early in embryogenesis, our cells somehow count how many X chromosomes they contain, then permanently inactivate all except one randomly selected X. At very early stages in development, both X chromosomes are active, but X-inactivation is initiated as cells begin to differentiate, occurring at the late blastula stage in mice, and most probably also in humans. Inactive X chromosomes remain in a highly condensed heterochromatic state throughout the cell cycle and can be seen as the Barr body (sex chromatin) on the periphery of the cell nucleus (see Figure 5.4).

The choice to inactivate the maternal or the paternal X is made randomly. But whichever of the parental X chromosomes is chosen for inactivation within a cell, that same X is inactivated in all daughter cells (**Figure 6.20A**). An adult female is thus a mosaic of cell clones, each clone retaining the pattern of X-inactivation that was established in its progenitor cell early in embryonic life. X-chromosome inactivation is strikingly revealed by

mixed coat colors in tortoiseshell and calico cats, both of which are almost always females (a small minority are XXY males). Tortoiseshell cats have a combination of two coat colors other than white (often orange and black); calico cats have additional white patches, as shown in Figure 6.20B.

X-inactivation is stable through mitosis but not across the generations. A woman's maternal X can equally well have been the active or inactive one in her mother and has the same chance as her paternal X of being inactivated in her own cells.

XIST RNA and initiation of X-inactivation

Inactivation of a human X chromosome is initiated at an X-inactivation center (XIC) at Xq13. It then propagates along the whole length of the chromosome in what may be an extreme example of the tendency of heterochromatin to spread (we consider heterochromatin spreading in more detail in the context of disease in Section 6.3).

The transient pairing of the two XIC sequences is probably the mechanism by which the X chromosomes are counted. Within this region, the *XIST* gene encodes a 17 kb spliced and polyadenylated noncoding RNA, an **X**-**i**nactivation-**s**pecific **t**ranscript expressed exclusively from the *inactive* X chromosome.

XIST is centrally involved in spreading heterochromatinization outward from the XIC: both *XIST* RNA and the Polycomb proteins it recruits seem to spread along the inactive X to initiate gene silencing along the length of the chromo-some. As a result, the inactive X has epigenetic marks typical of heterochromatin (H3K9me2, H3K9me3, H3K27me3, unmethylated H3K4, deacetylated H4, and frequent replacement of histone H2A by the macro-H2A histone variant). In differentiated cells that have already undergone X-inactivation, loss of *XIST* does not cause reactivation. That is, *XIST* is needed to establish X-inactivation but not to maintain it.

The mechanism of X-inactivation remains poorly understood. In addition to *XIST,* there are multiple other longer ncRNAs within the XIC. Several of them are known to have roles in the X-inactivation mechanism in mouse, but the organizations of the human XIC and its mouse counterpart, Xic, are rather different.

Escaping X-inactivation

A few genes on the X have active counterparts on the Y, notably in the terminal pseudoautosomal regions (but also in some other areas—see Figure 5.7). X-inactivation is therefore not a blanket inactivation of the entire chromosome, because no dosage compensation is needed for genes on the X that have functional equivalents on the Y. However, unlike in mouse, in which only a small number of genes escape X-inactivation and are not coated by *XIST* RNA, about 15 % of genes on the human X somehow escape inactivation.

6.3 ABNORMAL EPIGENETIC REGULATION IN MENDELIAN DISORDERS AND UNIPARENTAL DISOMY

Abnormal regulation of how our genes and other functional DNA sequences work can arise in different ways. Changes in DNA sequence may affect how DNA sequences work without necessarily causing any great change in their chromatin environment. In Chapter 7 we look at how disease arises directly as a result of altered base sequences and copy number variation. In this section we are largely concerned with abnormal epigenetic regulation. That occurs in rare cases when two copies of the same chromosome are abnormally inherited from just one parent. In addition, some genetic disorders show abnormal epigenetic regulation.

Principles of epigenetic dysregulation

An abnormal epigenetic change (**epimutation**) at one or more loci can be the immediate cause of pathogenesis in Mendelian disorders that show abnormal epigenetic regulation. However, the *primary* event is often a genetic mutation at a defined locus that may be one of several types. It may be a gene that makes a protein or RNA that

controls epigenetic modifications at other genes located elsewhere in the genome. It may be a *cis*-acting regulatory sequence that regulates epigenetic modifications of neighboring genes. In each case the epimutations determine the disease phenotype; because they lie downstream of a primary genetic event, they are often classified as *secondary epimutations* (**Figure 6.21**).

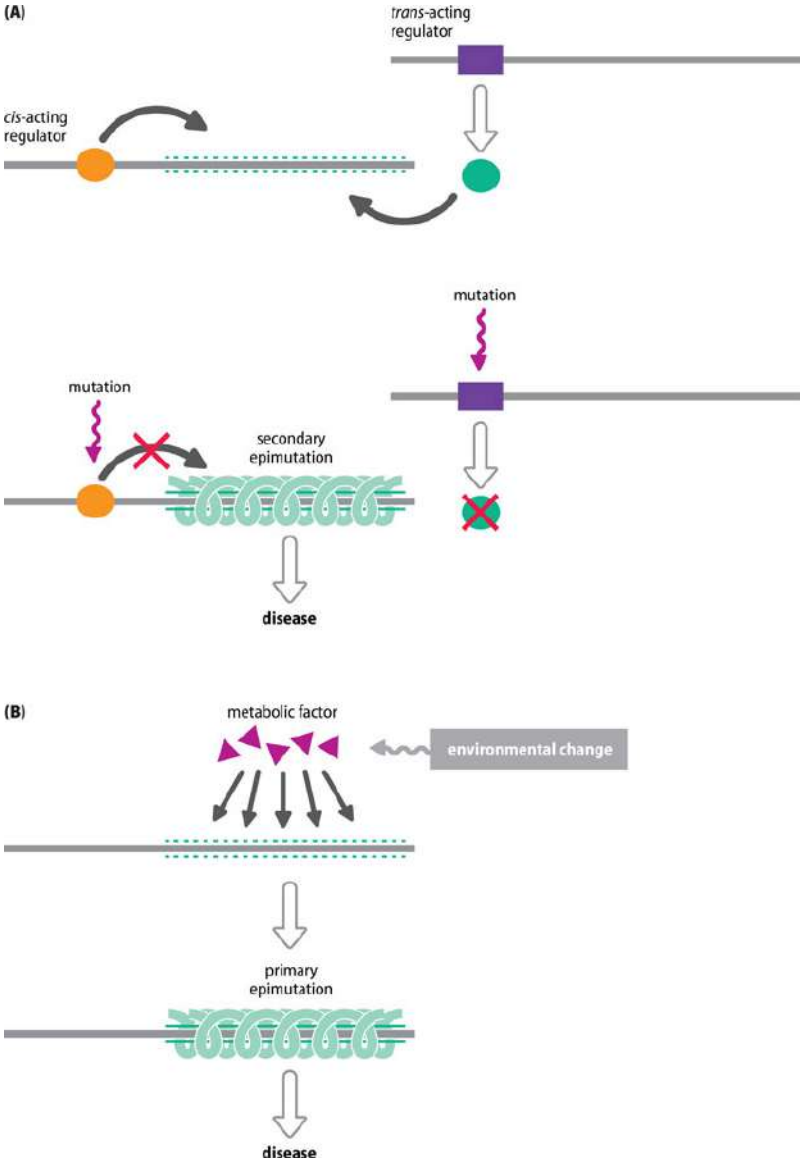


Figure 6.21 Primary and secondary epimutations. (A) Secondary epimutations arise through a change in chromatin state initiated by mutation at a *cis*-acting or *trans*-acting epigenetic regulator. Alterations to the regulators cause a change in chromatin state—in this case from a transcription-permissive environment (upper panel) to a repressive heterochromatic environment (lower panel). (B) Primary epimutations can effect a change in chromatin state without any change in the base sequence. Here, we imagine that an environmental change has changed the concentration of some metabolic factor that is important in DNA or chromatin modification, causing the change in chromatin state.

By contrast, *primary epimutations* can arise without any change to the base sequence. Instead, a chromatin state is reprogrammed, for example by some environmentally induced change, in a way that changes epigenetic controls (changes in metabolic factors may affect DNA methylation or histone modification states, and so on). Primary epimutations may be important in complex disease, and in Chapter 8 we examine the roles of both genetic and

epigenetic factors in complex disease. Epigenetic factors have a particularly important role in cancer, and we consider these separately in Chapter 10. Here we focus on epi-genetic dysregulation that arises through abnormal chromosome segregation or that is a feature of certain Mendelian disorders.

"Chromatin diseases" due to mutations in genes specifying chromatin modifiers

As detailed above, epigenetic marks are inscribed by a series of chromatin "writers", enzymes that methylate DNA or add different types of chemical group to defined amino acids on core histones. They can be recognized and bound by specific protein "readers" or be removed by specific enzymes ("erasers").

Individual genes that produce a chromatin writer, eraser, or reader can potentially regulate very many different genes across the genome, and a mutation in a gene of this type may result in heritable abnormal chromatin organization at multiple loci. In some cases the disruption to normal gene regulation might be incompatible with life because many of the target genes of chromatin modifiers are important in early development. Some disorders, however, do arise through mutations at chromatin modifier loci. These so-called "chromatin diseases" typically result in developmental disorders that can vary in phenotypes but are usually accompanied by mental retardation (**Table 6.7**). Affected individuals typically do not reproduce, usually presenting as sporadic (isolated) cases.

TABLE 6.7 EXAMPLES OF CHROMATIN DISEASES, DISORDERS THAT ARISE FROM MUTATION IN A CHROMATIN MODIFIER GENE

Class and type of chromatin modifier		Gene	Associated disease (reference)	Phenotype		
				Developmental	MR	Other
Writers	DNA methyl transferase	*DNMT3B*	ICF syndrome (OMIM 242860)	facial anomalies	variable	immunodeficiency; centromeric instability
	histone acetyltransferase	*CREBBP* or *EP300*	Rubinstein-Taybi syndrome (PMID 20301699; OMIM 180849)	characteristic facial features; digit anomalies	yes	
Erasers	histone lysine demethylase	*KDM5C*	Claes-Jensen type of syndromic, X-linked mental retardation (OMIM 300534)	variable—often mildly dysmorphic facial features; microcephaly	yes	
Readers	meCG-binding protein	*MECP2*	Rett syndrome (PMID 20301670; OMIM 312750)	see Clinical Box 2 on next page	variable	see Clinical Box 2 on next page
	chromatin remodeler	*ATRX*	α-thalassemia X-linked mental retardation syndrome (PMID 20301622; OMIM 301040)	cranial, facial, skeletal, and genital abnormalities; developmental delay and microcephaly	yes	thalassemia *(ATRX* regulates the α-globin genes among many others)

In **Clinical Box 2** we give a case study of a representative chromatin disease, Rett syndrome, an X-linked progressive neurodevelopmental disorder that results from mutations in the *MECP2* gene and affects girls almost exclusively. The regulatory protein MECP2 recognizes and binds 5-methylcytosine and is especially important in neuron maturation; the lack of a normal MECP2 protein in cells leads to an inability to recognize DNA methylation. The phenotype is partly determined by the X-inactivation pattern: inactivating the normal X in a relatively high proportion of neurons would be expected to result in a particularly severe phenotype. Failure to produce any functional MECP2 protein was initially expected to be lethal and would explain why affected males

are so rarely seen. (Affected males can occur as a result of a post-zygotic inactivating mutation or have certain missense *MECP2* mutations and severe neonatal encephalopathy.)

CLINICAL BOX 2 A CASE STUDY: RETT SYNDROME

Evie was referred to the Genetics Clinic when she was 30 months old because of delayed development and recent loss of skills. She was born by emergency Caesarian section because of meconium staining and abnormal CTG (cardiotocograph) trace (used for early detection of fetal distress in the third trimester), but she did not require resuscitation. She breast fed well but was always floppier and less interactive than her two older sisters. She sat unsupported at nine months but had a tendency to flop her head forward when sitting.

Her parents described a definite loss of skills from around two years of age. She started bottom shuffling at 18 months but by the age of two she had stopped attempting to move. She stopped babbling and saying "bye" and being able to hold toys or feed herself finger foods. She had episodes suggestive of absence seizures where she became unresponsive for several seconds. Around this time there was also a change in her behavior, and she interacted less with others and became increasingly bad-tempered. Evie showed stereotypic hand wringing movements and frequently brought her hands to her mouth and chewed them. She had generalized low muscle tone and was able to weight bear (with hyperextension at the knees), but not to crawl, or pull to stand. Her head circumference, on the 75th centile at birth, was now on the 9th centile, suggesting slowing of head growth.

Evie had already been extensively investigated for a metabolic cause for her developmental regression and had had a normal cranial MRI scan and EEG. Methylation studies for Angelman syndrome were normal. Her developmental regression and hand stereotypies (repetitive hand movements) were, however, suggestive of Rett Syndrome. *Rather than show a photo of Evie, we refer readers to a YouTube video describing another girl with Rett syndrome where the hand stereo- typies can clearly be seen: https://www.youtube.com/watch?v=H2iKz1Cx-HQ.*

To check if Evie had Rett syndrome, Sanger sequence analysis was carried out on the *MECP2* gene and revealed a heterozygous pathogenic variant c.455C>G p.(Pro152Arg). This variant has been reported many times in girls with Rett syndrome, and functional studies have shown it to affect MECP2 protein function. Evie's mother was tested and did not carry the variant; it was thus assumed to be *de novo*.

Evie was reviewed when she was four years old. She had not shown any further regression but had made very little developmental progress. She had developed three different types of seizure: nocturnal, tonic clonic, and absences, and was on two different anticonvulsant drugs. She had also developed a disordered breathing pattern, with periods of hyperventilation and breath-holding. She was only able to manage mashed food and had had several choking episodes.

Disease resulting from dysregulation of heterochromatin

Epigenetic regulation causes distinctive patterns of heterochromatin and euchromatin to form in our cells. Heterochromatin is first formed at nucleation sites, consisting of either repetitive DNA or silencer elements, and can then expand across long distances on a chromosome, even a whole chromosome, as in X-chromosome inactivation. Heterochromatin spreading involves converting open chromatin to condensed transcriptionally silent chromatin and is facilitated by communication between nucleosomes.

To avoid silencing essential genes, cells have evolved different mechanisms to limit the spread of heterochromatin. One such mechanism depends on **barrier elements**, a type of boundary element, that are able to protect genes from their surrounding environment. Barrier elements can include sequences that are selected to be comparatively nucleosome-free to provide a break in the nucleosome chain.

Altered heterochromatic states can impair normal gene expression in two quite different ways. Sometimes active genes are inappropriately exposed to heterochromatin controls and are silenced. An alternative form of dysregulation involves a reduction in heterochromatin and loss of gene silencing.

Inappropriate gene silencing

Aberrant heterochromatin regulation can silence genes that are normally meant to be expressed. A long-range **position effect** can mean that a gene is relocated to a position very close to constitutive heterochromatin (by a chromosome translocation or inversion, for example). In these cases, the boundary between euchromatin and heterochromatin can be reset, and the gene is silenced by heterochromatin spreading (**Figure 6.22**).

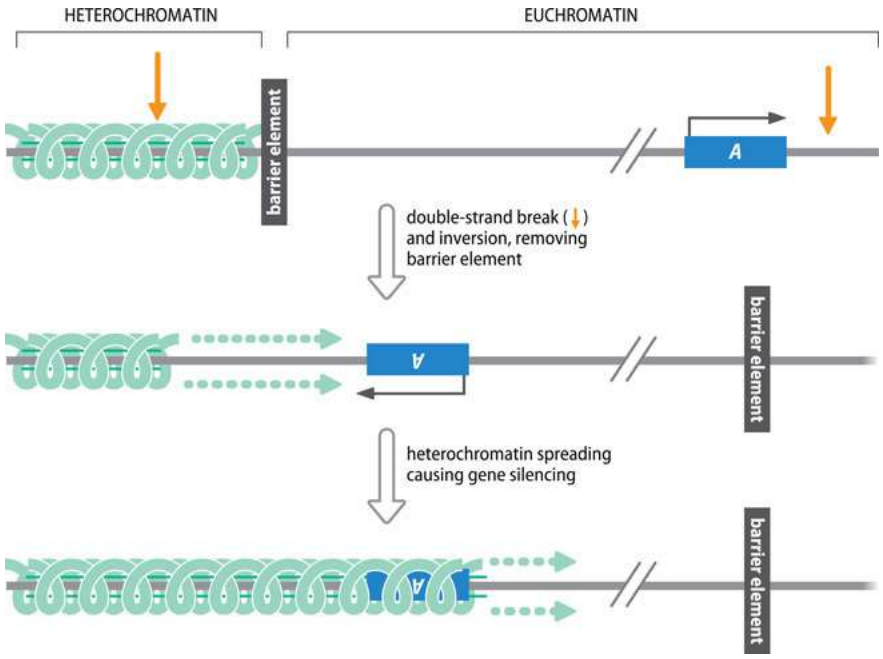


Figure 6.22 Heterochromatin spreading after an inversion causes displacement of a nearby barrier element. A *barrier element* protects genes in the euchromatin region from being silenced by adjacent heterochromatin (shown here in green). Large-scale rearrangements, such as the inversion shown here, can relocate the protective barrier element so that it no longer separates gene A from a neighboring heterochromatic region. This allows heterochromatinization of what had been a euchromatic region, resulting in gene silencing (a *position effect*).

Some special types of mutation can also induce heterochromatin formation within or close to a gene and so silence it. That can happen in the case of very large expansions of noncoding triplet repeats that occur in some recessively inherited disorders such as fragile X-linked mental retardation and Friedreich's ataxia. We consider this type of abnormal heterochromatin within the context of disease due to unstable oligonucleotide repeat expansion in Section 7.3.

Heterochromatin reduction

A primary function of some tumor suppressor genes such as *BRCA1* is to maintain the integrity of constitutive heterochromatin. As described in Chapter 10, mutations in these genes can result in a loss of heterochromatin organization; reduced centromeric heterochromatin leads to mitotic recombination and genome instability.

Some mutations causing heterochromatin reduction and inappropriate gene activation also result in inherited disorders. A classic example occurs in facioscapulohumeral dystrophy (FSHD), the third most common form of muscular dystrophy. This dominantly inherited disorder occurs as a result of simultaneous inheritance of two genetic variants:

- a reduction in heterochromatin due to unequal crossover at an array of tandem macrosatellite repeats at 4q35, close to the telomere, with each repeat containing a transcriptionally repressed *DUX4* retrogene copy

- a variant that creates a polyadenylation site close to the most telomeric of the repeats, enabling the most telomeric repeat to become transcriptionally active.

The combination of the two variants allows inappropriate expression of the DUX4 transcription factor, which is normally silenced in somatic cells and has been thought to be toxic to muscle cells—see **Figure 60.23A,B**.

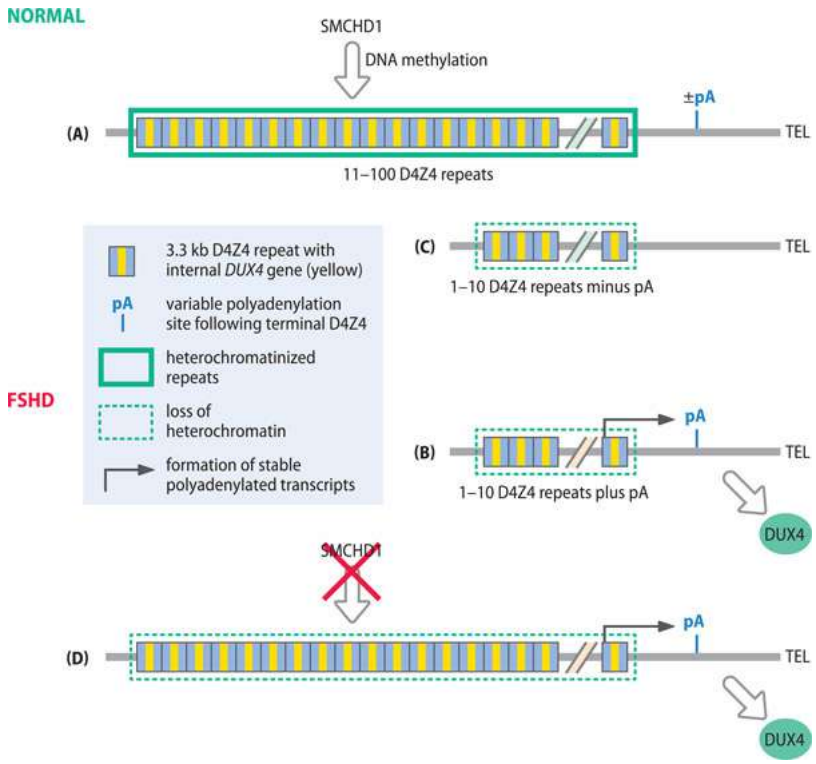


Figure 6.23 Heterochromatin reduction and inappropriate activation of the heterochromatic *DUX4* retrogene in facioscapulohumeral dystrophy (FSHD). (A) A normal chromosome 4 with a heterochromatinized array of 11–100 D4Z4 repeats, each ~3.3 kb long (the repeat copy number varies because of unequal crossover). A variable polyadenylation site is located adjacent to the last (most telomeric) D4Z4 repeat. (B) In FSHD1, the D4Z4 array is reduced in size to 1–10 repeats, causing a marked decrease in heterochromatin; the downstream polyadenylation site is present, allowing both transcription and translation of the last *DUX4* sequence. (C) In the absence of the downstream polyadenylation signal, the *DUX4* sequence cannot produce stable transcripts even when the heterochromatin is decreased. (D) In FSHD2 the downstream polyadenylation sequence is present, together with a long D4Z4 array that nevertheless has decreased heterochromatin because of a failure to produce the SMCHD1 methylation regulator.

One further complication is that the disorder is genetically heterogeneous: in the common FSHD1 form, illustrated by the case study shown in **Clinical Box 3**, the reduction in heterochromatin is caused by significant reduction in the number of D4Z4 repeats within the macrosatellite array at 4q35. In a second form, FSHD2, the reduction in heterochromatin is caused by mutation in the *SMCHD1* gene on chromosome 18 that regulates DNA methylation (Figure 6.23D).

CLINICAL BOX 3 A CLINICAL CASE STUDY: FACIOSCAPULOHUMERAL DYSTROPHY

John was born at term from an uneventful pregnancy to nonconsanguineous parents. He had normal development through childhood and enjoyed playing regular sports. As an adult, John continued to be active, attending the gym four times a week and regularly playing rugby. When he was 20 years old, John developed difficulty raising his right arm above shoulder height. The arm was not painful but appeared to him to be weak. He was initially

referred to an orthopedic surgeon, who passed him on to rheumatology. No cause was identified, and he was ultimately referred to the neuromuscular centre.

When first seen by the neuromuscular team at the age of 22 John had mild facial weakness, wasting of the pectoralis muscle on the right, and asymmetrical scapular winging, as shown in **Figure 1**. Shoulder abduction was asymmetrically weak, scoring 3- on the MRC score on the right, but 5- on the left. Other shoulder movements including adduction and internal and external rotation were slightly impaired. His lower limbs were stronger and almost completely normal.

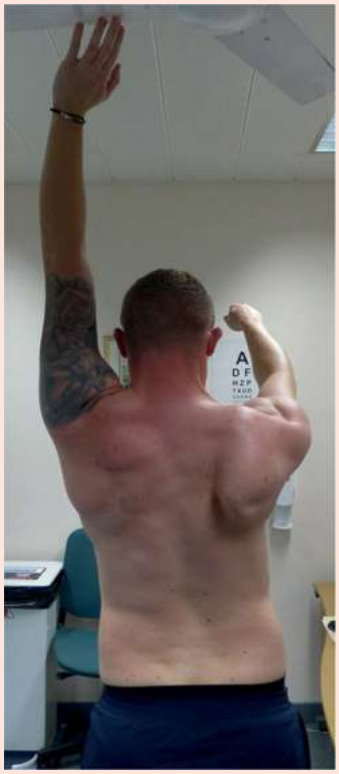


Figure 1 Asymmetrical scapular winging.

The neuromuscular consultant suspected a diagnosis of facioscapulohumeral muscular dystrophy and arranged testing for FSHD1. The D4Z4 repeated region at 4q35 can be investigated using a P13E11 probe on Southern blots of *Eco*RI-digested genomic DNA. When this was done, John's DNA sample revealed an abnormally small *Eco*RI fragment of 24 kb, suggesting that one of his chromosome 4s has only around six D4Z4 repeats and the contraction in the normal size of the D4Z4 repeat region has allowed the terminal *DUX4* gene to become, inappropriately, active. (The *DUX4* gene is strongly expressed in germ cells where the DUX4 protein works as a transcriptional activator but normally it is transcriptionally repressed in somatic cells, as a result of heterochromatinization.)

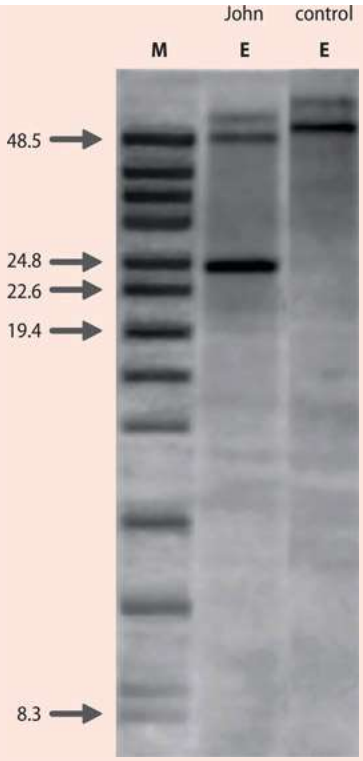


Figure 2 Southern blot analysis after pulsed field gel electrophoresis. Hybridization of a D4Z4 repeat region DNA probe, P13E11, to *Eco*RI-digested genomic DNA normally reveals a band of >40 kb, as shown in the unaffected control at right, but in John's DNA, the detected *EcoRI* fragment is only 24 kb. M, marker DNA lane with 13 size standards, five of which are identified by arrowheads. E, *Eco*RI-digested genomic DNA. Image kindly provided by Sarah Burton-Jones, South West Genomics Hub, Southmead Hospital, Bristol.

Uniparental disomy and disorders of imprinting

Recall that the sperm and egg genomes each carry epigenetic marks that are rather different from each other. Although the great majority of the gametic epigenetic marks are then erased in the early embryo, remaining imprints are retained in our somatic cells. We have more than 100 classically defined imprinted genes, and many of them have important roles in early development. In some cases, the maternal allele is consistently silenced or preferentially silenced (that is, monoallelic expression in some cell types, but biallelic expression in others); for other genes, it is the paternal allele that is consistently or preferentially silenced.

Occasional cases of uniparental diploidy (producing an androgenetic or gynogenetic embryo as shown in Figure 6.19) can occur. They are invariably lethal in the early embryo (each embryo of this kind fails to express multiple imprinted genes needed for fetal development). Sometimes, however, abnormal regulation of imprinted genes is confined to genes on a single chromosome and results in a developmental disorder. This can arise by a change in DNA sequence at or near the imprinted gene locus or by abnormal epigenetic regulation of the imprinted gene that may result as a downstream effect, often because of other genetic changes.

Uniparental disomy arises when a zygote develops in which both copies of one chromosome originated either from the father or from the mother. It occurs most often after a trisomic conceptus is first formed with two chromosome homologs from one parent and a single chromosome copy from the other parent; loss of the latter chromosome very early in development results in a heterodisomy (**Figure 6.24A**). The alternative is monosomy rescue, which results in isodisomy (with two identical copies of a chromosome—see Figure 6.24B). As described in the next section, uniparental disomy can result in a disorder of development if the chromosome happens to contain imprinted genes that are important in development.

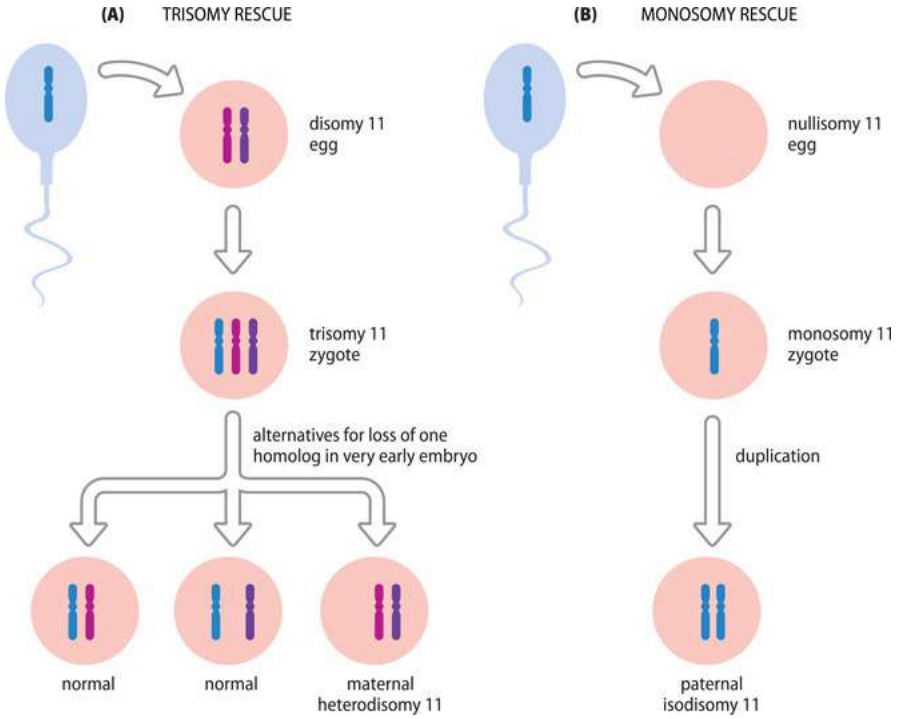


Figure 6.24 Uniparental disomy can arise by post-zygotic trisomy rescue or monosomy rescue. (A) Shown in the center is one type of trisomic zygote (with two maternal homologs plus one paternal chromosome, in this case chromosome 11). Trisomy 11 is lethal, but the trisomy can be corrected in the very early embryo by the loss of a chromosome 11 from one embryonic cell that then has a growth advantage and goes on ultimately to form an individual with the correct number of chromosomes. The disomic cell (and individual) may be normal (one paternal 11 plus one maternal 11) or have the two maternal chromosome 11 homologs (uniparental *heterodisomy*). (B) Monosomy rescue can occur by chromosome duplication, but in this case the result is uniparental *isodisomy* (the two chromosomes are identical, not homologs).

Abnormal gene regulation at imprinted loci

At an imprinted locus only one of the two parental alleles is consistently expressed (in at least some tissues), and alteration of the normal pattern of monoallelic expression can result in disease. Certain imprinted genes are important in fetal growth and development, so in these cases abnormal expression often results in recognizable developmental syndromes. Sometimes the normally expressed allele is not present or is defective, and deficiency of a gene product causes disease. In other cases, disease can be due to overexpression of a dosage-sensitive gene (Table 6.8). Analysis of human imprinting disorders has helped us understand the underlying gene regulation, and we provide here a background to two such imprinted gene clusters, as detailed below.

TABLE 6.8 EXAMPLES OF IMPRINTED DISORDERS OF EARLY CHILDHOOD

Disorder	Diagnostic clinical features	Molecular basis of imprinting disorder	Cause	
			UPD**	Others include
PATHOGENESIS DUE TO UNDEREXPRESSION OF GENES AT IMPRINTED LOCI				

* DD, developmental delay; IUGR, intrauterine growth retardation.

** UPD, uniparental disomy; pat., paternal; mat., maternal.

*** ICR1, ICR2, imprinting control regions 1 and 2.

Disorder	Diagnostic clinical features	Molecular basis of imprinting disorder	Cause	
			UPD**	Others include
Prader-Willi syndrome	DD*; low birth weight; hypotonia; hyperphagia	silencing/lack of active allele for imprinted genes at 15q11.2, including multiple *SNORD116 (HB11-85)* snoRNA genes (see Figure 6.26)	Mat.15, ~25%	Δpat.15q11-13,~70%
Angelman syndrome	DD (severe); no speech; epilepsy; ataxia	silencing/lack of active allele for imprinted *UBE3A* gene at 15q11.2 (see Figure 6.26)	Pat.** 15, ~5%	Δmat.15q11-13,~75%
Silver-Russell syndrome	IUGR*; faltering growth; short stature	silencing/lack of active allele for *IGF2* at 11p15.5 (see Figure 6.25) or for *MEST(PEG1)* at 7q31-32 *(IGF2* and MESTare maternally imprinted)	Mat.** 11	loss of pat. ICR1 methylation, ~35-50%
			Mat.7, ~8%	
PATHOGENESIS DUE TO OVEREXPRESSION OF GENES AT IMPRINTED LOCI AND/OR FETAL GROWTH PROMOTION				
Beckwith-Wiedemann syndrome	macrosomia/overgrowth; macroglossia; umbilical defect	biallelic expression of *IGF2* at 11 p15.5 (normally silenced on maternal 11) and/or biallelic expression of a ncRNAthat suppresses a growth-restricting gene, *CDKN1C;* see **Figure 6.22**)	Pat.11	loss of mat. ICR2*** methylation, ~50 %; gainofmat.ICR1*** methylation,~5%; *CDKN1C* mutation, ~5 %
Transient neonatal diabetes	IUGR*; neonatal diabetes with remission	biallelic expression of PLAGL1, a regulator of insulin secretion, at 6q24 (normally silenced on maternal chromosome 6)	Pat.6	30%

* DD, developmental delay; IUGR, intrauterine growth retardation.

** UPD, uniparental disomy; pat., paternal; mat., maternal.

*** ICR1, ICR2, imprinting control regions 1 and 2.

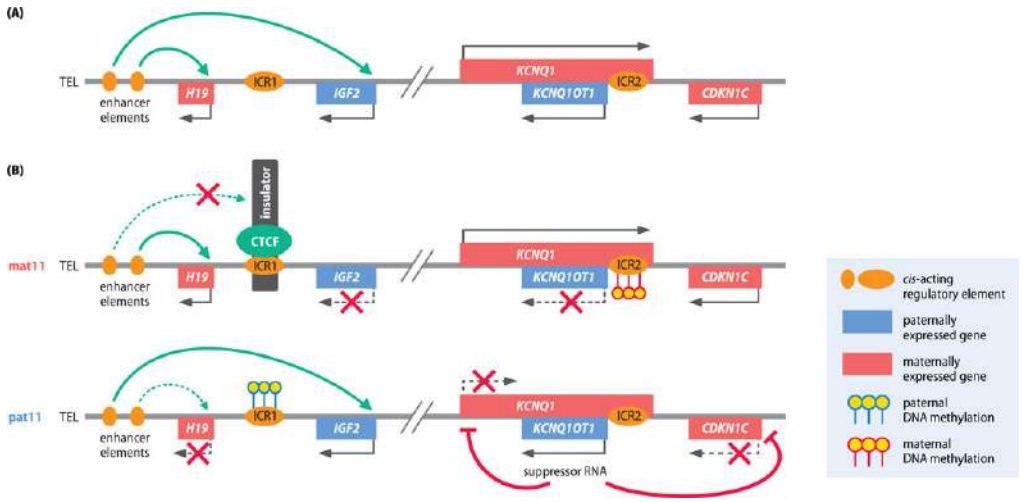


Figure 6.25 Imprinting control mechanisms in the 11p15 imprinted gene cluster. (A) Terminal genes and imprinting control regions (ICR1, ICT2) in the human 11p15 imprinted gene cluster. The ~600 kb gap (//) between the insulin-like growth factor 2 gene (*IGF2*) and

KCNQ1 contains at least five other imprinted genes. Arrows indicate the direction of transcription. TEL, telomere. (B) Regulation of imprinted genes in the 11p15 cluster on maternal and paternal chromosome 11 (mat11 and pat11). ICR1 acts as an insulator. It is hypomethylated on mat11 and bound by the CTCF protein, which recruits other proteins to act as a barrier, blocking enhancer elements close to *H19* from activating the distant *IGF2* gene. On pat11, ICR1 is extensively methylated and is not bound by CTCF, allowing the enhancer elements to preferentially activate *IGF2* instead of *H19*. ICR2 is located in an intron of *KCNQ1* and acts as a promoter for the antisense RNA gene *KCNQ1OT1,* which encodes a *cis*-acting suppressor RNA. ICR2 is hypomethylated on pat11, allowing transcription of the *KCNQ1OT1* suppressor RNA, which inhibits the transcription of the neighboring genes *KCNQ1OT1* and *CDKN1C*. On mat11, ICR2 is extensively methylated, blocking transcription of *KCNQ1OT1* and allowing the expression of the neighboring genes.

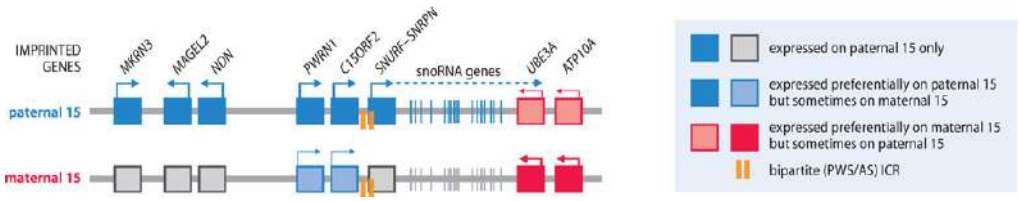


Figure 6.26 The imprinted gene cluster at 15q11–q12 associated with Prader-Willi syndrome (PWS) and Angelman syndrome (AS). Arrows show the direction of transcription of the indicated genes. The prominent long dashed blue arrow signifies that there is a long transcription unit with multiple noncoding exons that has been proposed to overlap the *UBE3A* gene. Numerous snoRNA genes (individual vertical blue lines) are found in introns of the long transcription unit, two of which are present in multiple copies: *SNORD116* (previously HBII-85) and *SNORD115* (previously HBII-52). A bipartite imprinting control region (ICR) is located near the promoter region of *SNURF-SNRPN*

The imprinted gene cluster at 11p15.5

A cluster of at least 10 imprinted genes in the subtelomeric 11p15.5 region has been well studied because of associations with Beckwith-Wiedeman syndrome (PMID 20301568) and many cases of Silver-Russell syndrome (PMID 20301499). The gene cluster has two different imprinting control regions, ICR1 and ICR2.

Figure 6.25A shows some key 11p15.5 genes regulated by ICR1 and ICR2. Both ICR1 and two nearby enhancer elements regulate *IGF2* (insulin growth factor type 2; paternally expressed) and *H19* (which makes a maternally expressed ncRNA). ICR2 regulates the *KCNQ1* gene (which makes a maternally expressed potassium channel), the *KCNQ1OT1* antisense RNA transcript (*KCNQ1* opposite strand transcript 1; paternally expressed), and *CDKN1C* (a suppressor of cell proliferation; maternally expressed).

ICR1 and ICR2 are each activated when hypomethylated and suppressed when extensively methylated. But they have opposite parental imprints: a maternally inherited chromosome 11 has a hypomethylated ICR1 and extensively methylated ICR2, but a paternal chromosome 11 has an extensively methylated ICR1 and a hypomethylated ICR2. They also use rather different control mechanisms (see Figure 6.25B).

Disease can result from significant changes in the methylation patterns or the base sequences of the ICRs and key imprinted genes. The most frequent cause of Silver-Russell syndrome is hypomethylation of ICR1 on both chromosome 11s so that both *IGF2* alleles are silenced. This can often happen by maternal 11 disomy or duplication of maternal 11p (see Table 6.8).

Beckwith-Wiedemann syndrome is marked by fetal overgrowth. It can occur when ICR1 is extensively methylated on both chromosome 11s so that both *IGF2* alleles are expressed, causing excessive growth. It can also occur when ICR2 is hypomethylated on both chromosome 11s, causing silencing of both alleles of the growth-restricting gene *CDKN1C*. Paternal disomy 11 is a common cause.

The imprinted gene cluster at 15q11-12

Another well-studied imprinted gene cluster located at 15q11–12, is associated with two neurodevelopmental disorders: Angelman syndrome and Prader-Willi syndrome. Angelman syndrome phenotypes include severe intellectual disability and microcephaly, and affected individuals are prone to frequent laughter and smiling. In Prader-Willi syndrome (PWS), affected individuals show mild intellectual disability and hyperphagia leading to obesity.

A single, bipartite imprinting control region regulates the whole cluster, which contains very many imprinted genes that are expressed only or preferentially on the paternal chromosome 15, including many small nucleolar RNA (snoRNA) genes, but just two imprinted genes that are preferentially expressed on the maternal chromosome 15 (**Figure 6.26**). The imprinted snoRNA genes seem to be located within introns of an extended transcription unit that includes a few exons making two proteins (SNURF and SNURPN) plus a large number of noncoding exons. By overlapping *UBE3A* and possibly *ATP10A,* the very long transcripts might silence these two genes on paternal chromosome 15.

Angelman syndrome and PWS are both caused by genetic deficiency. In the former case, the key problem is loss or inactivation of the maternal *UBE3A* allele, which makes a ubiquitin-protein ligase (both *UBE3A* alleles are normally expressed in most tissues, but in neurons only maternal *UBE3A* is active). The PWS phenotype, however, is attributable to the deficient expression of different genes normally expressed only on the paternal chromosome 15, including *NDN,* which regulates adipogenesis, and *SNORD116 / HBII-85* genes, which make a type of snoRNA that, in addition to its standard snoRNA role, also acts to regulate alternative splicing in some target genes.

Angelman syndrome and PWS are primarily due to large deletions that remove the same 5 Mb region of DNA including the genes shown in Figure 6.26, either from maternal 15q11-q12 (Angelman syndrome) or paternal 15q11-q12 (PWS). This region is flanked by low-copy-number repeat sequences that make it inherently prone to instability. However. some cases may result from point mutation—see **Clinical Box 4** for an example.

CLINICAL BOX 4 A CASE STUDY: ANGELMAN SYNDROME

Keira was referred to the Genetics Clinic by her Pediatrician at 20 months of age. She was her parents' first child and her mother, Marie, was 20 weeks pregnant at the time of referral. Born after a normal pregnancy and delivery, Keira had difficulty breast feeding and was therefore bottle-fed. She was a colicky baby who frequently vomited feeds and at three months was diagnosed as having gastroesophageal reflux. This resolved by one year of age. Her parents were concerned that her motor development was delayed by the time she was 6 months old. She started crawling at 12 months and by 18 months was able to pull to stand and cruise. At 20 months she was not yet pointing and had no single words. She was an extremely happy baby who was always smiling and laughing, but had occasional episodes, each lasting a few seconds, where she would go quiet and be unresponsive.

On examination, the geneticist noticed that she had jerky movements and walked holding hands with feet turned out and an unsteady wide based gait. She drooled copiously, laughed a lot and clapped her hands together frequently. She had a relatively small head circumference in comparison to height and weight and a wide mouth. The geneticist suspected that Keira had Angelman syndrome and requested a chromosomal microarray and methylation studies of the Angelman ICR. There was no evidence of a microdeletion of chromosome 15q11-13 or loss of the unmethylated (maternal) *SNRPN* allele. Targeted clinical exome analysis of a panel of 96 genes associated with Syndromic Intellectual Disability was performed and showed a heterozygous pathogenic SNV in the *UBE3A* gene: c.1749dup p.(Glu584Ter). Parental testing showed that Keira had inherited the pathogenic UBE3A variant from her mother, Marie (**Figure 1**).

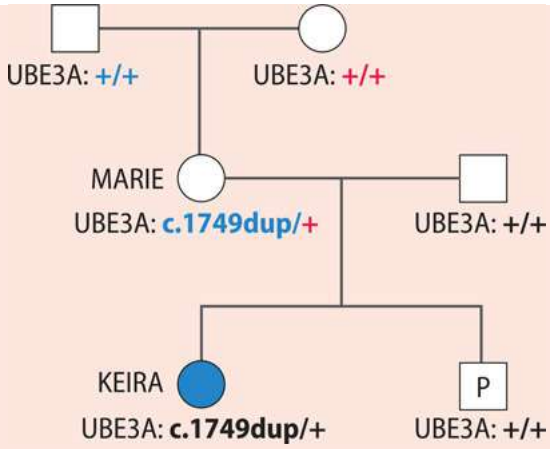


Figure 1

The pregnancy had therefore been at 50 % risk of having Angelman syndrome. Analysis of the cord blood of Keira's baby brother did not detect the pathogenic *UBE3A* variant and he has shown no clinical features suggestive of Angelman syndrome Marie has been offered prenatal diagnosis in a future pregnancy. Marie's parents both tested negative for the *UBE3A* variant. The *UBE3A* variant is a *de novo* variant in Marie, but she does not have Angelman syndrome, having inherited the variant from her father. *UBE3A* is paternally imprinted in the brain, and so the maternal *UBE3A* allele is active but the paternal allele is not expressed. The c.1749dup p.(Glu584Ter) variant is not uncommon—see **Figure 2** for an unrelated boy with Angelman syndrome carrying this variant *de novo*.

Figure 2

Imprinting and assisted reproduction

Another aspect of imprinting disorders relates to concerns about apparently increased frequencies of these disorders in births in which assisted reproductive technology (ART) has been employed. *In vitro* fertilization is now well accepted in economically advanced societies, where it accounts for 1–4 % of births. Because early embryogenesis is a critical time for epigenetic regulation and is sensitive to environmental factors, ART might impose added stress on embryos that can result in altered epigenetic profiles.

Imprinting disorders are very rare, and statistical support for increased incidence in assisted conception is difficult to achieve. But studies in mice have shown that although intracytoplasmic sperm injection does introduce

primary epimutations, they are normally corrected by epigenetic reprogramming in the germline and are therefore not transmitted to subsequent generations.

SUMMARY

- Genetic regulation of gene expression is dictated by the base sequence; epigenetic regulation of gene expression is independent of the base sequence.
- Epigenetic regulation of gene expression is rou tinely achieved by controlling chromatin structure. According to need, "open" chromatin structures form in some regions (allowing access to transcription factors); in other regions the chromatin is highly condensed, and transcriptionally inactive.
- *Cis*-acting regulatory sequences are located on the same *individual* DNA or RNA molecule as the sequences they regulate.
- *Trans*-acting gene regulators migrate in the cell to bind target sequences on DNA or RNA molecules. *Trans*acting regulatory proteins bind to targets by using nucleic acid-binding domains; *trans*-acting regulatory RNA molecules bind by base pairing.
- Gene promoters are composed of multiple short sequences. Core promoter elements are bound by ubiquitous transcription factors; other *cis*-acting regulatory elements are often bound by tissue-specific or developmental-stage-specific regulators.
- RNA splicing is largely dependent on recognition of *cis*-acting RNA sequence elements at splice junctions. Additional splice enhancer and splice suppressor sequences can be located in both introns and exons.
- Multiple different transcripts are produced for almost all of our genes and can give rise to alternative protein isoforms, increasing functional variation.
- Post-transcriptional regulation is often performed by microRNAs. An individual miRNA binds to partly complementary sequences in untranslated regions of multiple target mRNA sequences and downregulates expression.
- Open and condensed chromatin structure depends heavily on the extent of chemical modification of both DNA (via methylation of certain cytosines) and his-tones (notably by acetylation, methylation and phosphorylation of the side chains of certain amino acids).
- DNA methylation and histone modification patterns are prominent examples of epigenetic marks that can be stably inherited from one cell generation to the next. The appropriate chemical groups are added or removed by dedicated enzymes known, respectively, as "writers" or "erasers".
- Methylated CG dinucleotides and chemically modified amino acids on histones are bound and interpreted by specific proteins ("readers") that induce structural and functional changes in chromatin.
- ATP-dependent chromatin remodeling complexes can slide nucleosomes along the DNA to increase or decrease the spacing between nucleosomes, respectively allowing or denying access to transcription factors.
- *Cis*-acting long noncoding RNAs also work in epigen etic gene regulation in mammalian cells. They remain attached to the DNA strand from which they were transcribed and can serve as antisense RNAs (inhibiting expression of sense RNAs transcribed from the complementary DNA strand) or by forming scaffolds for binding certain *trans*-acting protein complexes.
- Genomic imprinting in mammals is an epigenetic phe nomenon whereby certain genes are either expressed or silenced according to whether they reside on a paternally transmitted chromosome, or a maternal one.

- X-inactivation means that one of the two X chro mosomes in women (and female mammals) is heterochromatinized.
- Barrier elements separate heterochromatin from neighboring euchromatin regions. If they are deleted or relocated by inversions or translocations, the neighboring euchromatin region can be heterochromatinized, causing gene silencing (a position effect).
- Uniparental disomy means that a pair of homologous chromosomes has been inherited from one parent. Disease can result when both chromosomes carry one or more imprinted genes.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

Enhancers, silencers, insulators, and general gene regulation

Ali T (2016) Insulators and domains of gene expression. *Curr Opin Genet Dev* 37:17–26; PMID 26802288.

Kolovos P (2012) Enhancers and silencers—an integrated and simple model for their function. *Epigen Chromatin* 5:1; PMID 22230046.

Latchman DS (2015) *Gene Control*, 2nd ed. Garland Science.

Alternative splicing and RNA editing

Kim E (2008) Alternative splicing: current perspectives. *BioEssays* 30:38–47; PMID 18081010.

Tang W (2012) Biological significance of RNA editing in cells. *Mol Biotechnol* 52:1–100; PMID 22271460.

MicroRNAs and competing endogenous RNAs

Baek D (2008) The impact of microRNAs on protein output. *Nature* 455:64–71; PMID 18668037.

Poliseno L (2010) A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 465:1033–1038; PMID 20577206.

Epigenetics in gene regulation (general)

Bonasio R (2010) Molecular signals of epigenetic states. *Science* 330:612–616; PMID 21030644.

Portela A & Esteller M (2010) Epigenetic modifications and human disease. *Nat Biotechnol* 28:1057–1068; PMID 20944598.

Histone modifications and DNA methylation

Chen Z & Riggs AD (2011) DNA methylation and demethylation in mammals. *J Biol Chem* 286:18347–18353; PMID 21454628.

Deaton AM & Bird A (2011) CpG islands and the regulation of transcription. *Genes Dev* 25:1010–1022; PMID 21576262.

Smallwood SA & Kelsey G (2012) De novo DNA methylation: a germ cell perspective. *Trends Genet* 28:33–42; PMID 22019337.

Suganuma T & Workman JL (2011) Signals and combinatorial functions of histone modifications. *Annu Rev Biochem* 80:474–499; PMID 21529160.

Long noncoding RNA in genetic and epigenetic regulation

Gil N & Ulitsky I (2020) Regulation of gene expression by *cis*-acting long non-coding RNAs. *Nature Rev Genet* 21:102–117; PMID 31729473.

Mercer TR & Mattick S (2013) Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol* 20:300–307; PMID 23463315.

Statello L (2021) Gene regulation by long non-coding RNA and its biological functions. *Nat Rev Mol Cell Biol* 22:96–118; PMID 33353982.

Genomic imprinting, X-inactivation, heterochromatin spreading

Barkess G & West AG (2012) Chromatin insulator elements: establishing barriers to set heterochromatin boundaries. *Epigenomics* 4:67–80; PMID 22332659.

Barlow DP & Batolomei MS (2014) Genomic imprinting in mammals. *Annu Rev Genet* 45:379–403; PMID 21942369.

Galupa R & Heard E. (2018) X-chromosome inactivation: a crossroads between chromosome architecture and gene regulation. *Annu Rev Genet* 52:535–566; PMID 30256677.

Haig D (2000) Kinship theory of genomic imprinting. *Annu Rev Ecol Syst* 31:9–32.

Epigenetic regulation and disease

Demars J & Gicquel C (2012) Epigenetic and genetic disturbance of the imprinted 11p15 region in Beckwith-Wiedemann and Silver-Russell syndromes. *Clin Genet* 81:350–361; PMID 22150955.

Hahn M (2010) Heterochromatin dysregulation in human diseases. *J Appl Physiol* 109:232–242; PMID 20360431.

Lemmers RJLF (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* 329:1650–1653; PMID 23143600.

Lemmers RJ (2012) Digenic inheritance of an *SCMD1* mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet* 44:1370–1374; PMID 20724583.

Sahoo T (2008) Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat Genet* 40:719–721; PMID 18500341.

Soellner L (2017) Recent advances in imprinting disorders. *Clin Genet* 91:3–13; PMID 27363536.

7

How genetic variation in DNA and chromosomes causes disease

DOI: 10.1201/9781003044406-7

CONTENTS

In Chapter 4, we outlined some basic principles of genetic variation. We covered the different types of genetic variation in our genome: both large-scale changes that make up structural variation and point mutations. And we related how sequence variation at the level of gene product relates to sequence variation at the level of DNA. Here we begin to consider the small fraction of human genetic variation that causes disease.

In this chapter we describe the different mechanisms that cause pathogenic DNA changes in genes and their consequences for the phenotype. The emphasis is on rare, highly penetrant variants that cause single-gene disorders, and on chromosome abnormalities. We do however take a broad view of protein dysregulation in disease. In Chapter 8 we go on to consider variants of low penetrance that confer susceptibility to common complex disorders, and in Chapter 10 we examine how genetic variation, predominantly somatic mutations, contributes to cancer.

In Section 7.1 we give an overview of how genetic variation results in disease. According to the number of nucleotides that are changed (and the number of genes affected), we will consider the pathogenic changes as occurring at different levels (different mutation mechanisms can be involved, depending on the size of the DNA change). We describe in Section 7.2 different kinds of pathogenic point mutations. They typically change just a single nucleotide (the most common mutation), or a very few nucleotides, and primarily affect the expression of a single gene.

As described in Sections 7.3 and 7.4, various genetic mechanisms also produce moderate-to large-scale mutations (affecting from tens of nucleotides to a few megabases of DNA); they are often triggered by inappropriate pairing of repetitive DNA sequences. Some very large-scale mutations and chromosome breaks produce recognizable changes at the chromosome level as detected by light microscopy. They will be covered in Section 7.5 together with other types of chromosome abnormality resulting from errors in chromosome segregation and recombination.

As initially considered in Chapter 5, the link between deleterious mutations and disease phenotype is often not straightforward. A deleterious mutation that might be expected to result in a single-gene disorder, according to the principles described in Chapter 5, may produce different degrees of disease severity, or no disease at all. We examine factors that complicate the link between genetic variation and disease phenotype, beginning in Section 7.6 with the tiny mitochondrial genome: it has just 37 genes but is nevertheless important in disease, and has unique properties that

complicate the molecular pathology. Then in Section 7.7 we go on to describe factors affecting phenotypes that result from variation in the nuclear genome.

In Section 7.8 we examine downstream effects at the protein structure level, notably how altered protein folding, and protein aggregation contribute to disease phenotypes. Finally, in Section 7.9 we consider the difficulties in correlating genotypes and phenotypes, and we give examples of how the phenotype of a monogenic disorder can be influenced by various factors including genetic variation at other gene loci and environmental factors.

7.1 AN OVERVIEW OF HOW GENETIC VARIATION RESULTS IN DISEASE

The great majority of variation in our DNA appears to be without consequence. For the most part, that happens because just a small percentage of our genome is functionally important (the great majority of nucleotides within introns and in extragenic DNA can be changed by small mutations without any obvious effect on the phenotype). A second, and minor, reason is genetic redundancy: some genes are present in multiple, almost identical copies—an inactivating mutation in a single ribosomal RNA gene in nuclear DNA has no effect because each type of cytoplasmic rRNA is made by hundreds of extremely similar gene copies.

Pathogenic mutations do not occur haphazardly in our DNA. For example, single nucleotide substitutions, the most common type of pathogenic mutation, are not random: certain types of DNA sequence are more vulnerable to point mutation (mutation *hotspots*). And, as detailed below, different arrangements of repetitive DNA also predispose to different classes of mutation, including many large-scale DNA changes. The genetic variation that causes disease may do so by causing two broad changes at the level of gene product, as listed below.

- A change in the ***sequence*** of the gene product. The result may be a total loss of function—the mutant gene product is not produced, or is incapable of carrying out its normal task. Or it may have a significantly reduced ability to work normally (as a result of a hypomorphic mutation). Sometimes, it may acquire an altered function causing it to inhibit the working of a normal gene product produced by the other parental allele. Rarely, mutation causes the gene product to have a new function that is harmful in some way (causing

cells to die or to behave inappropriately). As we will see, loss of function and gain of function quite often involve a change in protein structure.

- A change in the ***amount*** of the gene product. This can happen in three major ways, as listed below.

    a. *Change in gene copy number*. Whole gene deletion or duplication can result from different mechanisms, including short-range unequal crossover (Section 7.4), and also errors in chromosome segregation and recombination (Section 7.5). In addition, gene amplification (see **Figure 7.1**) is common in many cancers, as explained in Chapter 10.

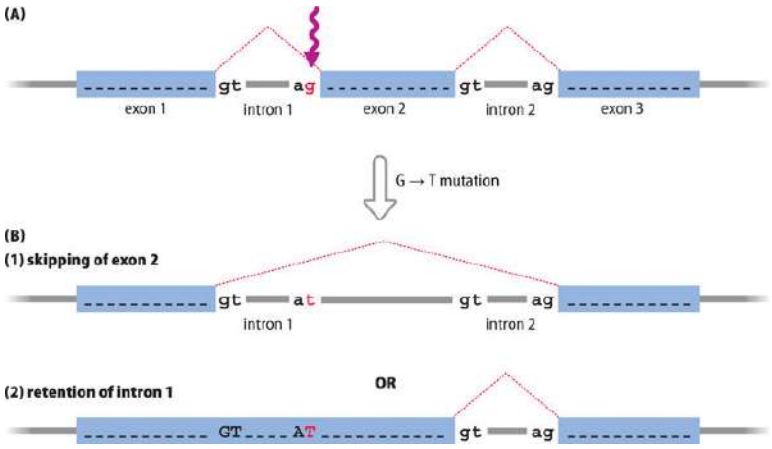


Figure 7.3 Splice-site mutations can cause exon skipping or intron retention. Exon sequences enclosed within blue boxes are represented by dashed lines, with upper-case letters indicating key individual nucleotides. Intron sequences are represented as gray horizontal lines with lower-case letters indicating key nucleotides. Dashed red lines show the positions where splicing will occur later *at the RNA level* to bring together transcribed exon sequences within RNA transcripts. (A) A normal situation with exons separated by introns with conserved 5¢ GT and 3¢ AG terminal dinucleotides. (B) Alternative outcomes of a G ® T mutation at the conserved 3¢ terminal nucleotide of intron 1, inactivating the splice acceptor site. Outcome (1): *exon skipping*. Using the next available splice acceptor site (at the 3¢ end of intron 2) causes skipping of exon 2 and a frameshift if the number of nucleotides in exon 2 is not a multiple of three. Outcome (2): *intron retention.* Alternatively, splicing of intron 1 is abandoned and the intron 1 sequence is retained in the mRNA, forming part of a large exon that also contains the original exon 1 and exon 2 sequences.

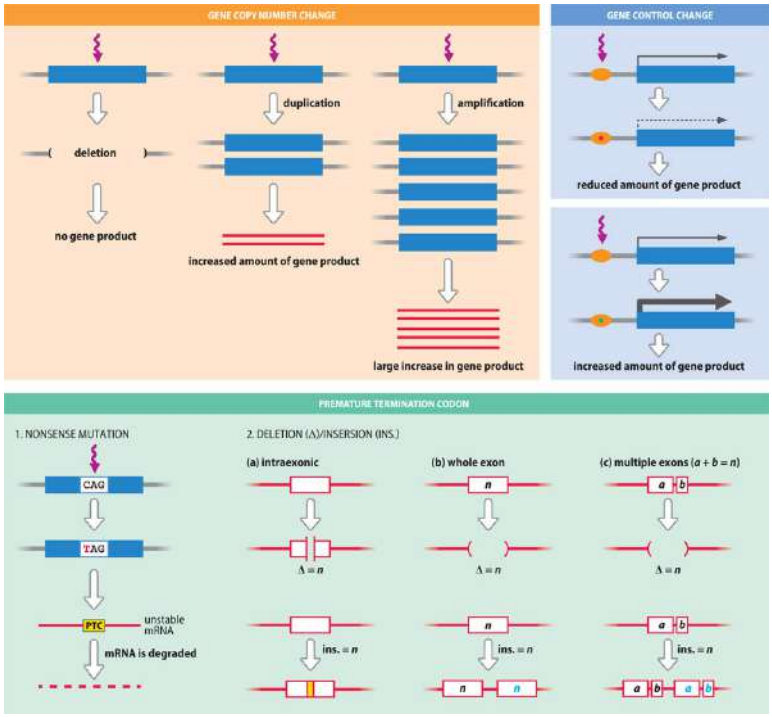


Figure 7.1 Classes of mutations that cause disease by altering the amount of a gene product. Mutations leading to premature termination codons (PTCs) frequently activate a pathway that leads to mRNA destruction (as explained in Box 7.1), and failure to make a protein. They are usually caused by nonsense mutations or by small or large frameshifting deletions and insertions (note: large insertions are often caused by duplications of one or more exons). If we denote the total number of inserted or deleted nucleotides of coding DNA as *n*, then a frameshift occurs when *n* is not exactly divisible by three. Occasionally (not shown here), a PTC may occur as a by-product of mutation that causes certain types of aberrant RNA splicing (exon skipping or intron retention, as described in Figure 7.3).

b. *Change in gene regulation.* Gene expression can be very significantly impacted by mutations affecting important upstream or intragenic regulatory sequences (or by epigenetic changes such as imprinting and position effects as described in Sections 6.2 and 6.3)
c. *Premature termination codon.* Premature termination codons often result in unstable mRNA and no protein product. They arise principally from two classes of mutation: nonsense mutations and frame-shifting insertions and deletions (Figure 7.1 gives an overview).

Although most pathogenic mutations affect individual genes, some mutations (and chromosome abnormalities) can *simultaneously* affect multiple genes. Large-scale deletions and duplications, for example, result in a simultaneous change in the copy number of multiple genes, with adverse effects. Additionally, mutations in some genes that produce *trans*-acting regulators can indirectly have consequences for multiple different target genes that they regulate.

The importance of repeat sequences in triggering pathogenesis

As detailed in Section 2.5, our genome has many types of repetitive DNA sequences, and various types of DNA duplication have been important in shaping the genome to allow the development of biological complexity. But despite the evolutionary advantage they confer, repeat sequences have a downside: they often predispose DNA molecules to undergo inappropriate pairing and subsequent sequence changes resulting in disease, and some repeats can cause disease by spontaneously and unpredictably inserting into the genome.

Pairing of DNA sequences in proper register is important at two different levels: (i) the paired individual strands of each DNA double helix and (ii) the paired double-stranded DNAs of chromatids of the same chromosome. *Tandem repeats*—those that are neighbours in a head to tail fashion (® ® ®)—can cause misalignment of the two single strands of a DNA helix and of the two double-stranded DNAs of paired chromatids. Resulting misaligned DNA strands in a double helix can undergo small deletions and duplications; misaligned chromatids can trigger moderate to large-scale deletions and duplications.

Interspersed repeats that may reside on the same nuclear DNA or mtDNA molecule, or on different DNA molecules, can also pair inappropriately with other repeats of the same type, predisposing to abnormal sequence exchanges. And certain families of interspersed *transposon repeat* can make copies of themselves that are able to insert elsewhere in the genome, including into genes. We cover the different mechanisms in later sections, but **Table 7.1** gives an overview.

TABLE 7.1 AN OVERVIEW OF HOW DNA REPEATS FREQUENTLY PREDISPOSE TO PATHOGENESIS

DNA repeats		Mechanisms (major mechanisms in italics)	Examples
Class	**Involvement**		

DNA repeats		Mechanisms (major mechanisms in italics)	Examples
Class	**Involvement**		
Tandem repeats	Short tandem repeats stabilize local mispairing of the two DNA	*Replication slippage* (DNA replication after slipped strand mispairing, causing strands to have fewer or more repeats)	Simple short tandem repeat variation causing insertions or deletions (Section 7.3).
	strands of a single DNA helix *(slipped strand mispairing)*	Mechanism poorly understood but involves slipped strand mispairing and DNA repair	Unstable short tandem repeat expansion (Section 7.3)
	Large tandem repeats stabilize local mispairing of two sister chromatids or two non-sister chromatids of a chromosome (Section 7.4).	Mispairing of large tandem repeats on opposing chromatids predisposes to: *unequal crossover (UEC); unequal sister chromatid exchange (UESCE);* and *gene conversion,* thereby producing deletions, duplications, or mutant genes	99% of cases of 21-OH deficiency are due to sequence exchanges occurring between mispaired large repeats containing the 21-hydroxylase gene and a related pseudogene (Clinical Box 6)
Interspersed repeats	Abnormal pairing of interspersed repeats in nuclear DNA	*UEC, UESCE,* or *intrachromatid recombination* causing deletions, duplications or inversions	Inversion in *F8* gene causing hemophilia A (Figure 7.11)
	Insertional inactivation by transposon repeat	*Retrotransposition* (cDNAof retro-transposon repeat inserts into genome)	

DNA repeats		Mechanisms (major mechanisms in italics)	Examples
Class	Involvement		
	Abnormal pairing of interspersed repeats in mtDNA (Section 7.6)	Mispairing of almost identical short repeats in mtDNA followed by cleavage and rejoining of fragments	Deletions in mtDNA (Table 7.13)

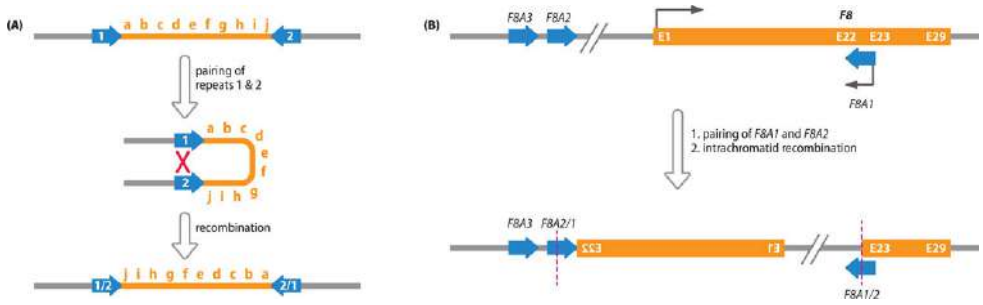


Figure 7.11 Intrachromatid recombination between inverted repeats produces inversions and is a common cause of hemophilia A. (A) Inverted repeats 1 and 2 on the same DNA strand can mispair by inducing looping of the intervening DNA. Subsequent recombination within the paired repeats produces hybrid repeat sequences (1/2 and 2/1) and inversion of the intervening DNA. (B) In about 50 % of cases with hemophilia A, the mutation is a large inversion that disrupts the blood clotting factor VIII gene *(F8)*. The 191 kb *F8* gene has 29 exons, and within the large intron 22 is a small gene, *F8A1,* that is transcribed from the opposite strand. *F8A1* is a member of a family of low-copy-number repeats that includes two closely related sequences, *F8A2* and *F8A3,* located upstream of *F8*. Mispairing between either of these repeats and *F8A1* can induce an inversion, by looping out of the intervening DNA to allow recombination between the mismatched repeats (such as *F8A1* and *F8A2,* as shown here; the dashed red vertical lines mark the boundaries of the inversion). The resulting inversion disrupts the *F8* gene, splitting it into two oppositely oriented fragments, one containing exons 1–22 and the other from exons 23–29.

7.2 PATHOGENIC NUCLEOTIDE SUBSTITUTIONS AND TINY INSERTIONS AND DELETIONS

Pathogenic single nucleotide substitutions within coding sequences

A single nucleotide substitution within a coding sequence has the effect of replacing one codon in the mRNA by another codon. There is, however, substantial

redundancy in the universal genetic code: as explained below, all amino acids other than methionine and tryptophan are specified by multiple codons (from two to six).

As a result of the redundancy in the genetic code, a mutated codon quite often specifies the same amino acid as the original codon. A coding sequence substitution such as this—one that does not change an amino acid—is known as a **synonymous substitution** (sometimes called a **silent mutation**). Because there is no change in amino acid, no change in phenotype might be expected. Nevertheless, as discussed below, a minority of synonymous substitutions nevertheless cause disease, almost always by simultaneously altering RNA splicing.

The alternative is a **nonsynonymous substitution**. There are different classes (Table 7.2), but the predominant one causes one amino acid to be replaced by another, a **missense mutation**. Missense mutations sometimes have minimal effects on the phenotype, but some of them have adverse effects as described below.

TABLE 7.2 CLASSES OF NONSYNONYMOUS MUTATION

Class	Definition	Example and comments
Missense mutation	an amino-acid-specifying codon is replaced by a codon for a different amino acid	GGA (glycine) is replaced by CGA (arginine).The effect is greatest when the replacement amino acid has very different physiochemical properties
Nonsense (stop-gain) mutation	an amino acid-specifying codon is replaced by a premature stop codon	a G → T substitution may result in codon GGA (glycine) being replaced by UGA (stop). Results in unstable mRNA or production of a truncated protein (see Box 7.1)

* The 3′ untranslated region will have termination codons in all three reading frames. As a result, in the case of a stop-loss mutation the "read-through" past the normal stop codon usually picks up another termination codon quite quickly, so that often the extended C-terminus is not long. The effect on protein function may often not be large unless the extended C-terminus causes problems for protein folding or protein stability.

Class	Definition	Example and comments
Stop-loss mutation	a natural stop codon is replaced by an amino acid-specifying codon	a T → G substitution may result in UGA (stop) being replaced by GGA (glycine), with translational *read-through* (the first part of the 3′ untranslated region is translated and the protein has an extended C-terminus*

* The 3′ untranslated region will have termination codons in all three reading frames. As a result, in the case of a stop-loss mutation the "read-through" past the normal stop codon usually picks up another termination codon quite quickly, so that often the extended C-terminus is not long. The effect on protein function may often not be large unless the extended C-terminus causes problems for protein folding or protein stability.

Relative frequencies of silent and amino-acid-replacing substitutions

The relative frequencies of single nucleotide substitutions that are silent and those that cause an amino acid to be replaced (missense mutations) vary according to the base position in a codon. If we first consider the genes in nuclear DNA, 61 codons can specify an amino acid (**Figure 7.2**). If we take an average, two out of every three substitutions at the third base position in a codon are silent; by contrast, 100 % of substitutions at the second base position and 184 out of 192 (about 96 %) of substitutions of the first base are nonsynonymous.

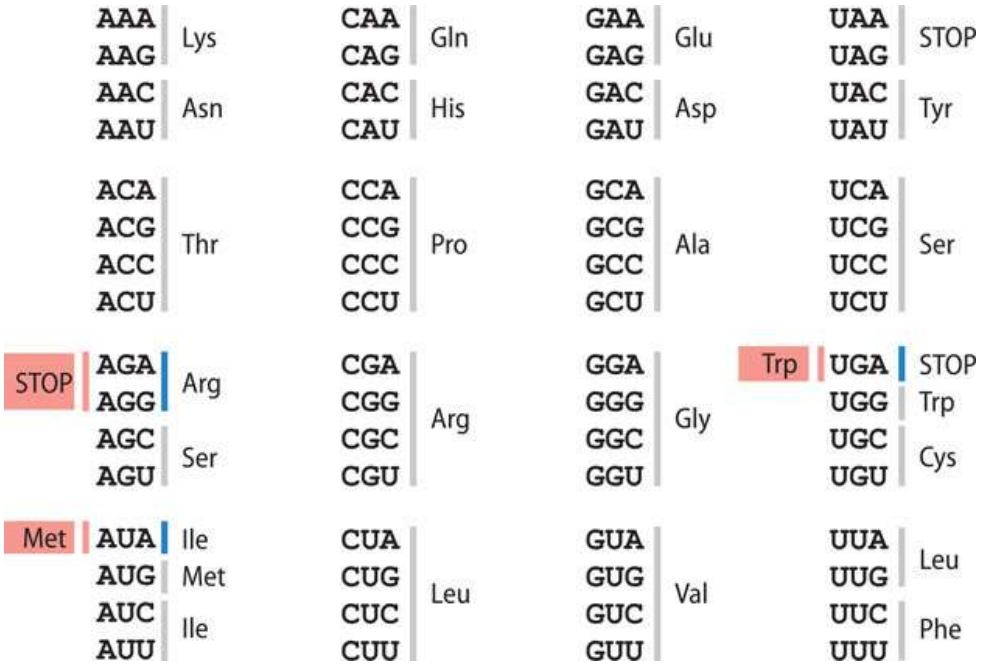


Figure 7.2 **The genetic code**. Pale gray bars to the right of codons identify 60 codons interpreted in the same way for nuclear and mitochondrial mRNA. Four codons—AGA, AGG, AUA, and UGA—are interpreted differently. Flanking blue bars and lettering to the right show the interpretation for nuclear genes; pale pink bars on the left show the interpretation for genes in mtDNA. The "universal" genetic code (for nuclear genes) has 61 codons specifying 20 different amino acids, with different levels of redundancy: from unique codons (Met, Trp) to sixfold redundancy (Arg, Leu, Ser). The remaining three codons—UAA, UAG, and UGA—normally act as stop codons (however, UGA can occasionally specify a 21st amino acid, selenocysteine, and UAG can occasionally specify glutamine). For genes in mtDNA, 60 codons specify an amino acid, and there are four stop codons (AGA, AGG, UAA, and UAG).

Consider, for example, a G ® A substitution that results in replacement of codon GG**G** by GG**A**. The genetic code shows that both the original GGG codon and the replacement GGA codon specify the amino acid glycine. If the substitution had been G ® C or G ® T (to give GG**C** or GG**U** codons), the altered codon again would have specified glycine. Like glycine, several other amino acids are exclusively, or largely, determined by the first two bases of a codon—there is flexibility in how the third base of a codon pairs with the 5¢ base of the anticodon (*base wobble*).

Genetic redundancy at the first base position is responsible for silent substitutions in some arginine and leucine codons. Thus, codons **A**GA and **A**GG specify arginine, as do codons **C**GA and **C**GG (an A ® C or C ® A change at the first base position is silent in these cases). Similarly, codons **C**UA and **C**UG specify leucine, as do codons **U**UA and **U**UG.

Conservative substitution: replacing an amino acid by a similar one

Nucleotide substitutions that change an amino acid can have different effects, according to the degree to which the replacement amino acid differs from the original amino acid (based on properties such as polarity, molecular volume, and chemical composition—see below). Perhaps fewer than 30 % of substitutions have no, or very little, functional significance; the remainder are roughly equally split into those with weak to moderate negative effects on protein function, and those with strongly negative effects.

A nucleotide substitution that replaces one amino acid by another of the same chemical class is a **conservative substitution** and often has minimal consequences for how the protein functions. **Table 7.3** lists different chemical classes of amino

acids plus some distinguishing features of individual amino acids (see Figure 2.2 for the chemical structures).

TABLE 7.3 AMINO ACIDS CAN BE GROUPED INTO SIX CLASSES ACCORDING TO THE CHEMICAL PROPERTIES OF THEIR SIDE CHAINS

Common feature of side chain		Amino acids*	Comments
Polar	Basic (positively charged)	Arg (R); Lys (K); His (H)	arginine and lysine have simple side chains with an amino ion ($-NH_3^+$); histidine has a more complex side chain with a positively charged imido group
	Acidic (negatively charged)	Asp (D); Glu (E)	simple side chains ending with a carboxyl ion (-COO-)
	Amide group	Asn (N); Gln (Q)	simple side chains ending with a $-CONH_2$ group
	Hydroxyl group	Ser(S);Thr(T);Tyr(Y)	serine and threonine have short simple side chains with a hydroxyl group; tyrosine has an aromatic ring
	Polar with sulfhydryl (-SH) group	Cys(C)	disulfide bridges (-S-S-) can form between *certain* distantly spaced cysteines in a polypeptide and are important in protein folding
Nonpolar		Gly (G); Ala (A); Val V); Leu (L); Ile (I); Pro (P); Met (M); Phe (F);Trp (W)	glycine has the simplest possible side chain—a single hydrogen atom. Phenylalanine and tryptophan have complex aromatic side chains

* See Figure 2.2 for structures of amino acids.

Nonconservative substitutions: effects on the polypeptide/protein

Replacing one amino acid by another belonging to a different chemical class may be expected to have more significant consequences. A key factor is whether the

individual amino acid has an important role in the function of the protein. It might play a vital role at the activation site of an enzyme, for example, be a critical part of a specific recognition sequence used to bind some interacting molecule, or have a side chain that needs to be chemically modified in a specific way for the protein to be functionally active.

Additional factors include the potential effects on protein folding and protein structure (for a brief summary of protein structure, see Section 2.1 and Box 2.2). Thus, for example, for thermodynamic reasons, globular proteins usually fold so that nonpolar, uncharged amino acids are buried in the interior and polar amino acids are on the outside, exposed to what is usually a hydrophilic aqueous environment; substitutions that change this pattern may induce incorrect protein folding.

Some amino acids are not tolerated in certain structural elements. Thus, for example, proline cannot be accommodated in an a-helix: if an amino acid is substituted by a proline the a-helix is disrupted. Conversely, certain amino acids have specific structural roles. Glycine (with the smallest possible side chain—a single hydrogen atom), and proline (the only amino acid in which the side chain loops back to rejoin the polypeptide backbone) are important in allowing the polypeptide backbone to bend sharply. They often have important roles in protein folding. The triple-stranded helical structure of collagens, for example, requires glycine at about every third residue; prolines (and hydroxyprolines) are also extremely frequent in collagens.

Cysteine has a unique role in protein folding. The sulfhydryl (−SH) groups on *certain* distantly located cysteines on the same polypeptide may interact to form a disulphide bridge (−S−S−); this can be important in establishing globular domains (such as for the immunoglobulin superfamily proteins in Figure 4.10). Replacing either cysteine by any other amino acid breaks the intrachain disulphide bond; as a result, cysteine is the most conserved amino acid in protein evolution.

In addition to causing simple loss of normal function or incorrect protein folding, missense mutations can also result in some new protein property that is damaging to cells and tissues in some way, or alters their behavior. We consider this aspect in more detail when we discuss the effects of genetic variants in Section 7.7.

Mutations that result in premature termination codons

A natural termination (stop) codon in mRNA triggers the ribosome to dissociate from the mRNA, releasing a polypeptide. However, many pathogenic mutations in

coding DNA cause an in-frame *premature* termination codon to be inserted into a coding sequence, either directly or indirectly.

Nonsense mutations are nonsynonymous substitutions that directly replace an amino-acid-specifying codon by a stop codon. For nuclear DNA, that means a substitution that produces one of three stop codons UAA, UAG, or UGA in the corresponding mRNA. Note that the genetic code for mitochondrial DNA is different —as shown in Figure 7.2.

A **frameshift mutation** may indirectly lead to a premature termination codon. Deletion or insertion of a sequence of *n* nucleotides in coding DNA produces a shift in the translational reading frame when *n*/3 *is* not an integer (Figure 1 in Box 2.1 on page 26 gives the principle). If a different reading frame is used, an in-frame premature termination codon is quickly encountered. Frameshifts often involve deletions or insertions at the DNA level, but they may also result from mutations producing altered splicing (exon skipping, intron retention), as described below.

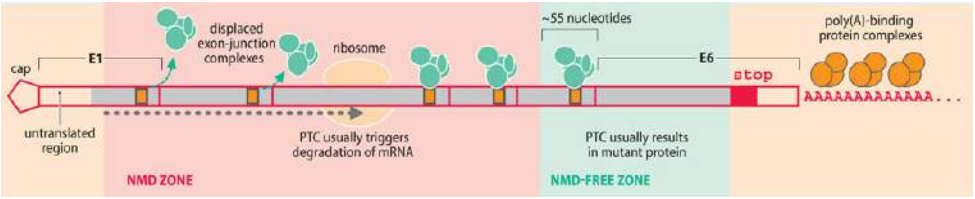


Figure 1 Nonsense-mediated decay. In mammalian cells, the primary nonsense-mediated decay (NMD) pathway is splicing-dependent. Certain components of the splicing machinery, called exon-junction complexes (EJCs), bind about 20–24 nucleotides upstream of the 3¢ end of each transcribed exon sequence and remain bound to the mature RNA. Here we illustrate a mature mRNA (with coding sequence in gray) formed from six exons (separated by thin vertical red lines; the first and last exons are labelled E1 and E6). Vertical orange boxes superimposed on the gray coding sequence show EJC-binding sites. The first ribosome to bind and then move along the mRNA will displace each EJC in turn until it reaches the stop codon and disengages. A premature termination codon (PTC) occurring up to 55 nucleotides before the end of the last exon (large pink box) leaves the mRNA with usually several EJCs attached, which normally triggers destruction of the mRNA. However, a PTC that occurs late in the mRNA, in the last exon and up to 55 nucleotides before it (large green box) usually means that the mRNA is translated to make a truncated protein that may sometimes give rise to a stronger phenotype than obtained by mRNA degradation.

At the DNA level, deletions or insertions of one or two nucleotides in coding DNA are a quite common cause of disease. As described in Section 7.3, short tandem repeats with mononucleotide or dinucleotides predispose to 1- to 2-bp insertions and deletions. In addition, intragenic deletions that remove one or more

exons or exon duplications often cause frameshifts (Figure 7.1). And transposons can occasionally accidentally insert into coding DNA (we consider large deletions and insertions like these in Section 7.4).

The usual result of nonsense mutations and frameshifting mutations is that the mRNA is degraded by a mechanism known as nonsense-mediated decay (**Box 7.1**). Occasionally, however, translation occurs to give a truncated protein. Truncated proteins produced in this way can sometimes interfere with wild-type proteins produced from the normal allele and so affect their function (a *dominant-negative effect)*. We consider the implications in Section 7.7.

BOX 7.1 NONSENSE-MEDIATED DECAY AS AN mRNA SURVEILLANCE MECHANISM

Various RNA surveillance mechanisms monitor RNA integrity, checking for splicing accidents (such as when transcribed intron sequences are inappropriately retained in the mRNA—see Figure 7.3B[2]) and occasional errors in base incorporation during transcription. These errors frequently give rise to in-frame premature termination codons (PTCs) that could be dangerous—the aberrant transcripts could give rise to truncated proteins that might have the potential to interfere with the function of normal proteins.

To protect cells, an mRNA surveillance mechanism known as **nonsense-mediated decay** (**NMD**) degrades most mRNA transcripts that have an in-frame PTC. The primary NMD pathway is dependent on RNA splicing (single exon genes escape NMD because they do not undergo RNA splicing). Multisubunit protein complexes, called exon-junction complexes, are deposited shortly before the 3¢ end of each transcribed exon during RNA splicing and remain bound at positions close to the exon-exon boundaries in mature RNA (**Figure 1**).

The first ribosome to bind and move along the mRNA displaces each of these complexes in turn before disengaging from the mRNA at the natural stop codon. If there is an in-frame PTC, however, the ribosome detaches from the mRNA at an early stage; some exon junction complexes remain bound to the RNA, which usually signals mRNA destruction. However, in-frame PTCs within or just before the last exon often escape NMD and are translated to give truncated proteins (Figure 1).

Nonsense mutations, frameshifting insertions and deletions, and certain splicing mutations (such as those that result in retention of an intron) can activate nonsense-

mediated decay.

Pathogenic splicing mutations

Many disease-causing mutations affect RNA splicing. The great bulk of them are in DNA sequences specifying *cis*-acting RNA elements that regulate how a specific gene undergoes RNA splicing, and this will be the focus here. Note, however, that disease can occasionally be caused by mutations in genes encoding *trans*-acting regulators of splicing.

As illustrated in Figure 6.4A, fundamental *cis*-acting regulatory elements that control RNA splicing are located at or close to splice junctions. Point mutations in these sequences often have marked effects on RNA splicing, especially if they change highly conserved nucleotides such as the GT (GU in RNA) and AG dinucleotides at the extreme 5¢ end and 3¢ end, respectively, of an intron. That can result in abnormal splicing patterns such as omission of an exon (**exon skipping**), or failure to splice out an intron *(intron retention*)—see **Figure 7.3**.

Pathogenic mutations can also occur in additional *cis*-acting splice regulatory elements, including splice enhancer and splice silencer sequences in exons and introns (see Figure 6.4B). Mutations like this may be less readily identified as pathogenic mutations and can explain why some mutations causing synonymous substitutions are pathogenic (**Figure 7.4A**).

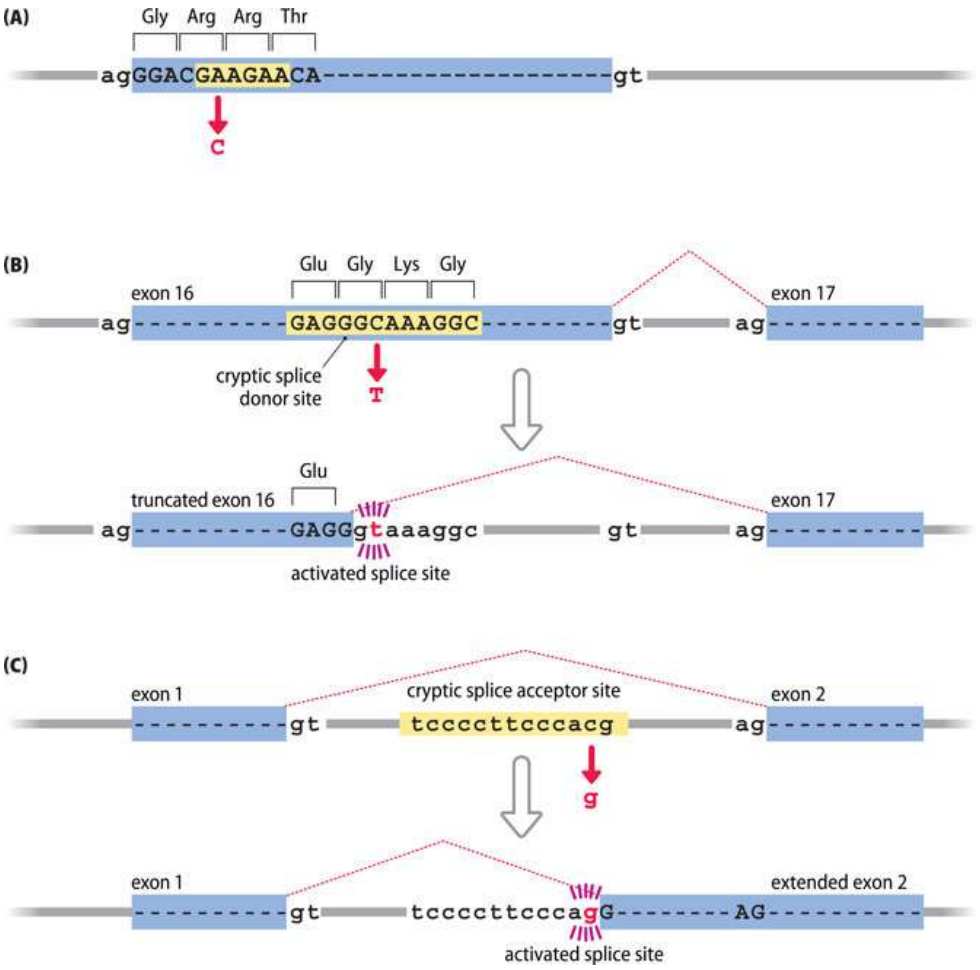


Figure 7.4 Apparently harmless synonymous and intronic substitutions can be pathogenic. Within exons (blue boxes) significant nucleotides are shown in capital letters; important intronic nucleotides are in lower case. Dashed red lines indicate splicing of exon sequences (occurring *at the RNA level*). (A) The A ® C mutation leads to replacing one arginine-specifying triplet (CGA) by another (CGC), but causes disease by changing an exonic splice enhancer sequence (highlighted in yellow) at the beginning of the exon. (B) A homozygous synonymous C ® T substitution in exon 16 of the calpain 3 gene replaces one glycine-specifying triplet GGC) by another (GGT); see PMID 7670461. It causes limb girdle muscular dystrophy by simultaneously activating a cryptic splice site (GAGGGCAAAGGC), to become a functional splice donor. As a result, exon 16 is truncated (the final 44 nucleotides are not included in RNA transcripts). The resulting shift in the translational reading frame produces a premature termination codon early in exon 17. (C) An apparently harmless single-nucleotide intronic substitution causes disease by activating a cryptic splice site closely resembling a 3¢ splice site (*splice acceptor*; the mutation results in the terminal AG dinucleotide required in splice acceptor sequences). Aberrant RNA splicing causes a sequence from the 3¢ end of intron 1 to be included in exon 2, extending it (causing disease by introducing a frameshift or by disturbing protein structure/protein folding).

By chance, sometimes a sequence may be almost identical to a genuine splice donor or splice acceptor site (a *latent* or **cryptic splice site**), and changing a single nucleotide can cause it to become a novel splice site. Activation of cryptic splice sites produces truncated or extended exons (see Figure 7.4B,C for examples).

Abnormal splicing can have variable consequences. For a protein-coding gene, a loss of coding exon sequences (by exon skipping or exon truncation) or a gain of coding exon sequences (by exon extension or intron retention) may result in a frameshift in the translational reading frame at the RNA level. In that case, the introduction of a premature termination codon might induce RNA degradation or produce a truncated protein (as described in Box 7.1).

If there is no shift in the translational reading frame, pathogenesis may nevertheless occur because of a loss of key amino acids, or, for exon extension, by the inclusion of extra amino acids that might destabilize a protein or impede its function. Intron retention in coding sequences would be expected to introduce a nonsense mutation (because of the comparatively high frequency of termination codons in all three reading frames).

Genesis and frequency of pathogenic point mutations

As detailed in Section 4.1, single nucleotide substitutions often arise as a result of spontaneous chemical degradation of DNA that has not been repaired effectively. Some types of single nucleotide substitution are especially frequent in human DNA. C ® T transitions are particularly common because the cytosine in CG dinucleotides is a hotspot for mutation in human (and vertebrate) cells. That happens because the CG dinucleotide is a target for cytosine methylation, and methylated cytosines tend to be deaminated to give thymines (as previously shown in Figure 4.3), which are not easily identified as altered bases by DNA repair systems.

Very small insertions and deletions are often produced by **replication slippage**, an error that typically occurs in DNA replication when a single nucleotide or short oligonucleotide is tandemly repeated. During DNA replication the nascent strand occasionally mispairs with the parent DNA strand (the mispairing is stabilized by base pairing between misaligned repeats on the two strands; the result is that the DNA polymerase either stutters at a tandem repeat or skips forward—see Figure 4.6 on page 91, for the mechanism). Arrays with multiple tandem repeats are particularly susceptible to replication slippage; simply by chance, coding sequences occasionally have sequences with such repeats. For example, in about one out of four occasions, on average, two consecutive lysines in a protein are specified by the hexanucleotide

AAAAAA. Any run of consecutive nucleotides of the same type means a significantly increased chance of replication slippage—in this case the daughter strands are liable to have five or seven A's, causing a frameshifting deletion or insertion.

Mutation rates in the human genome

Comprehensive genome sequencing in family members indicates that the genomewide germ line nucleotide substitution rate is 10^{-8} per nucleotide per generation. That equates to about 30 *de novo* nucleotide substitutions on average in the 3 Gb haploid genomes inherited from each parent.

The mutation frequency varies across chromosomes and genes. Some gene-associated features make them more likely mutation targets. Genes are GC-rich and have a higher content of the CG dinucleotide; the cytosine in CG dinucleo-tides is a mutational hotspot, undergoing C ® T transitions at a rate that is more than 10 times the background mutation rate.

The mitochondrial genome is extremely gene-rich, and the mutation rate is many times higher than in the nuclear genome. The mitochondrial genome is vulnerable, possibly because the great majority of reactive oxygen species are produced in mitochondria. Close proximity to these dangerous radicals results in much more frequent damage to the DNA, which is devoid of a protective chromatin coating, unlike nuclear DNA. Unrepaired DNA replication errors can also be significant in mtDNA.

Total pathogenic load

Only a small fraction of the novel changes that arise in the genomes we inherit from our parents is likely to be pathogenic, but because our parents are carriers of previously generated mutations our genomes contain many deleterious mutations. As yet there is no easy way to identify the total pathogenic load—all pathogenic mutations—within a genome.

Population genomics projects have shown that, depending on our ethnic background, each of us carries about 100 mutations that would be expected to result in loss of gene function (with an average of 20 genes that are homozygously inactivated), plus about 60 missense variants that severely damage protein structure.

One prediction is that the average person might have over 400 damaging DNA variants. That might seem an impossibly high load of pathogenic mutations but many of these mutations are common variants in non-essential genes, such as the *ABO* blood group gene (which is homozygously inactivated in people with blood group O).

Effect of parental age and parental sex on germ line mutation rates

Increased parental age often correlates with increased frequency of genetic disorders. We consider the maternal age effect in trisomy 21 in Section 7.5. For small-scale mutations there is often a higher frequency of *de novo* mutation in the male germ line, and paternal age effects can be apparent, as described below.

The frequency of *de novo* mutation can be expected to be high in gametes that have undergone many cell divisions since originating from the zygote through a **primordial germ cell** (the cells that are set aside in the early embryo to give rise to the germ line). That happens because DNA replication precedes each cell division, and mutations often arise as a result of uncorrected errors in DNA replication. Two meiotic cell divisions are required to form oocytes and sperm cells, but the number of preceding mitotic cell divisions required to produce the first meiotic cells is very different between the two sexes. All the egg cells that will be available to a woman are formed before birth. By contrast, after the onset of puberty in men, sperm are continuously being formed by the division of spermatogonial stem cells. The number of cell divisions required to produce gametes is therefore higher in men, and especially so in older fathers (**Figure 7.5**).

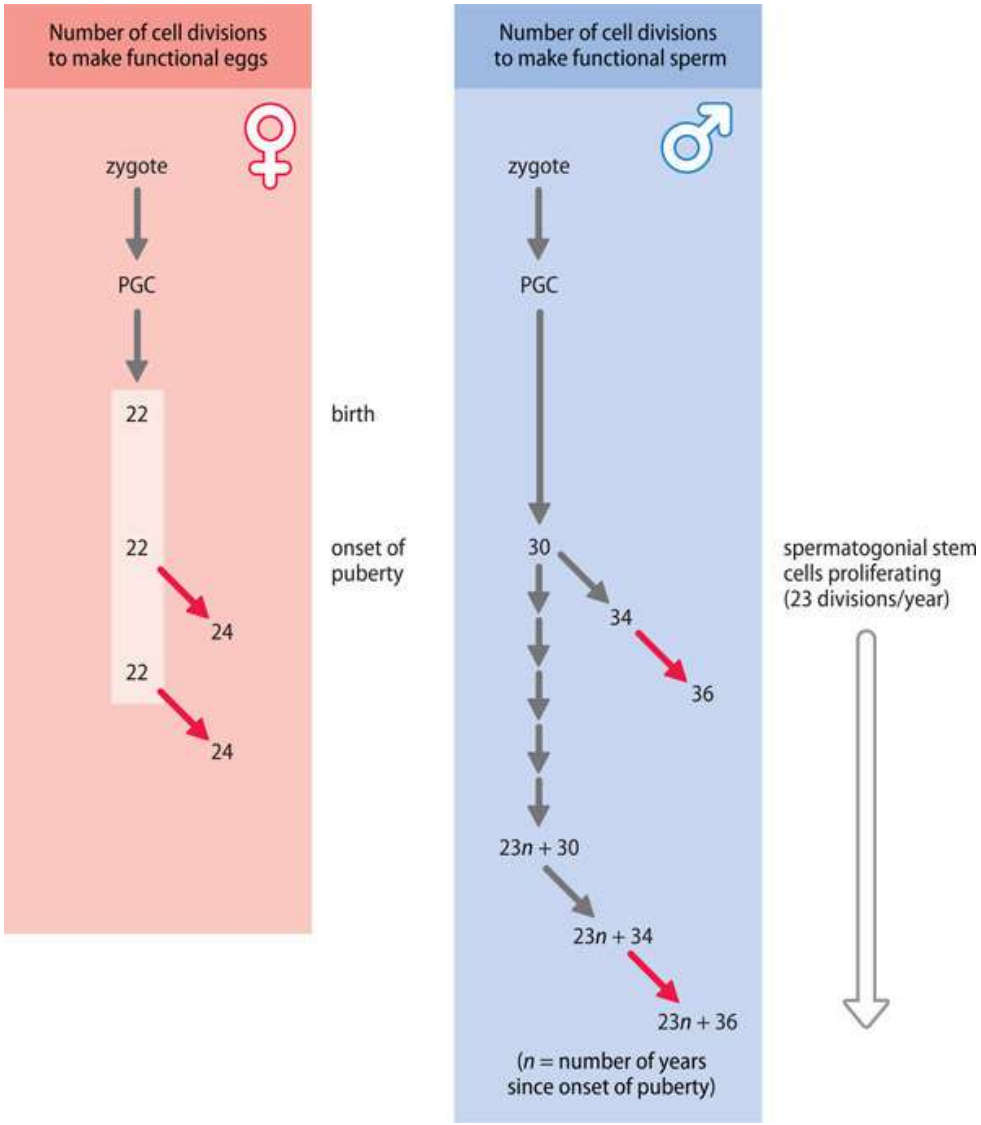


Figure 7.5 Sex differences and paternal age differences in the number of cell divisions required to make gametes. Numbers represent completed cell divisions en route from a human zygote to gametes. Sperm and egg cells are formed by two sequential meiosis (red arrows) preceded by multiple mitoses (gray arrows). In females, all the ~22 mitoses required to get to the first meiotic cell are accomplished before birth (part of meiosis I has been completed by then, but then suspended until activated by ovulation). No matter how old mothers are, a total of ~24 cell divisions separate zygote from egg cells. Males are different: gametogenesis continues throughout adult life. About 30 cell divisions separate zygote from the spermatogonial stem cell used to make the first sperm cells at the onset of puberty. From spermatogonial stem cell to gamete takes four mitoses and then two meioses. So, sperm cells produced at the onset of male puberty have gone through 30 + 4 + 2 = 36 cell divisions. Thereafter, spermatogonial stem cells divide about every 16 days (or about 23 times per year). If we take an average age of 14 years, say, for the onset of male puberty, sperm

from a 64-year-old man have been produced in a process requiring 36 + (23 × [64 − 14]) mitoses plus two meioses, or a total of 1186 cell divisions. PGC, primordial germ cell.

Paternal-age-effect disorders and selfish spermatogonial selection

A small group of exceptional congenital disorders occur spontaneously at remarkably high apparent rates, reaching 1 in 30 000 births for achondroplasia, with marked paternal age effects and paternal germ line transmission (see **Table 7.4** for some examples). That might suggest exceptional germ line mutation rates (up to 1000-fold higher than the average rate). However, studies of mutation rates in sperm have suggested that the underlying mutations (which are mis-sense mutations that change a single amino acid) do not occur at especially high frequencies. Instead, the mutations have been thought to result in mutant proteins that cause the dysregulation of spermatogonial stem cell behavior, thereby conferring a selective growth advantage on any spermatogonial stem cell that contains them. Stem cells containing these mutations might proliferate to reach high frequencies, explaining the paternal origin of many mutations and the paternal age effect. In each case the mutant proteins are fibroblast growth factor receptors or other proteins that work in the growth factor receptor-RAS signal transduction pathway. The underlying mutations belong to a class of gain-of-function mutations described in Section 7.7.

TABLE 7.4 EXAMPLES OF PATERNAL-AGE-EFFECT DISORDERS SUGGESTED TO BE ASSOCIATED WITH SELFISH SPERMATOGONIAL SELECTION

Disorder	Gene	Mutation/amino acid change	Estimated birth prevalence for new mutations	Parental origin of mutation
Apert syndrome	FGFR2	p.Ser252Trp p.Pro253Arg	~1/65000	100% paternal
Crouzon/Pfeiffer syndrome	FGFR2	>50 mutations	~1/50 000 to 1/100000	100% paternal
Achondroplasia	FGFR3	p.Gly380Arg*	1/30000	100% paternal

* The G380R change in the FGFR3 (fibroblast growth factor receptor 3) protein in achondroplasia is caused by either a G → C or a G → A change at nucleotide number 1138 in the reference cDNA sequence. For further information seethe paper by Goriely &Wilkie (2012) in Further Reading.

Disorder	Gene	Mutation/amino acid change	Estimated birth prevalence for new mutations	Parental origin of mutation
Muenke syndrome	FGFR3	p.Pro250Arg	~1/30000	100% paternal
Noonan syndrome	PTPN11	many mutations	~1/10000	100% paternal

* The G380R change in the FGFR3 (fibroblast growth factor receptor 3) protein in achondroplasia is caused by either a G → C or a G → A change at nucleotide number 1138 in the reference cDNA sequence. For further information seethe paper by Goriely &Wilkie (2012) in Further Reading.

Surveying and curating point mutations that cause disease

Point mutations are the most frequent contributors to disease, and data on known or suspected pathogenic point mutations have been curated in a variety of different databases. We consider how point mutations are identified as being pathogenic in Chapter 11, when we consider diagnostic DNA approaches. For now, we give some brief points in the subsections below.

Point mutations in coding DNA

The vast majority of pathogenic point mutations have been recorded in protein-coding genes. For coding DNA, identifying some types of disease-causing mutation is comparatively easy. Nonsense mutations and insertions or deletions that change a known translational reading frame almost scream “Pick me!” at the investigator.

Making the correct call for missense mutations and small non-frameshifting deletions or insertions is harder. Evolutionary and population genetic studies can often help here: substitution or deletion of an amino acid is much more likely to be pathogenic if that specific amino acid has been strongly conserved during evolution. As detailed in Chapter 11, various computer programs can be used. One important application is to assess the likelihood of a missense mutation being pathogenic on the basis of predicted differences in the physicochemical properties of the original amino acid and the substituted one. Querying databases of previously recorded mutations can also be very useful.

Point mutations in RNA genes and other noncoding DNA

Although we have significantly more RNA genes than protein-coding genes, causative mutations in monogenic disorders have been identified almost exclusively in protein-coding genes (see **Table 7.5** for some examples of the very few RNA genes that have been implicated). Explanations for this anomaly may include the general difficulty in identifying pathogenic mutations in noncoding DNA (which is generally poorly evolutionarily conserved and lacking a reading frame), and genetic redundancy for some genes (such as rRNA and tRNA genes in nuclear DNA).

TABLE 7.5 EXAMPLES OF RNA GENES MUTATED IN SINGLE-GENE DISORDERS

RNA class	Locus*	Disorder	PMID
miRNA	*MIR96*	Autosomal deafness, type 50 (OMIM 613074)	19363479
	MIR184	EDICT syndrome (OMIM 614303)	2199627S
	MIR204	Retinal dystrophy with ocular colomba (OMIM 616722)	26056285
snRNA	*RNU4ATAC*	microcephalic osteodysplastic primordial dwarfismtypei (OMIM 210710)	21474760
long noncoding RNA	*TERC*	Type 1 Autosomal dominant dyskeratosis congenital (OMIM 127550)	11574891
MT-rRNA	*MT-RNR1*	Inherited nonsyndromic hearing loss	7689389
MT-tRNA	Many	Various–covered in Section 7.6	33655490

* Interested readers can find recent reviews at PMID 24007299 and 32741549. PMID, PubMed Identifier. MT, mitochondrial.

Arguably, the most important explanation is that proteins are the major functional endpoints of cells (even though there are impressive layers of gene regulation by noncoding RNAs). Many RNA genes have not been well studied, but more recent studies have emphasized the importance of various noncoding RNAs in the pathogenesis of other disorders, such as cancers.

Databases of human pathogenic mutations

Human mutation databases range from large general databases to more specific ones (**Table 7.6**). Locus-specific databases focus on a specific gene (or sometimes genes) associated with an individual disorder. The submitted data include both pathogenic mutations and normal variants, and so the databases can be of help in evaluating whether newly identified mutations are likely to be pathogenic, as described in Chapter 11.

TABLE 7.6 EXAMPLES OF DIFFERENT TYPES OF DATABASES THAT CURATE DISEASE-ASSOCIATED MUTATIONS

Database	Description	Website
GENOMEWIDE DATABASES		
Human Gene Mutation Database	comprehensive data on germ line mutations in nuclear genes associated with human inherited disease	http://www.hgmd.cf.ac.uk/ac/index.php
COSMIC	comprehensive catalog of somatic mutations in cancer	https://cancer.sanger.ac.uk/cosmic
MITOMAP	mitochondrial genome database	https://www.mitomap.org/MITOMAP
LOCUS-SPECIFIC DATABASES (see also http://www.hgvs.org/locus-specific-mutation-databases and PMID 21540879)		
CFTR2	specific *CFTR* cystic fibrosis gene variants	https://cftr2.org/
MUTATION CATEGORY DATABASES		
SpliceDisease database	disease-associated splicing mutations	http://cmbi.bjmu.edu.cn/sdisease

The Human Genome Variation Society also maintains links to many other useful mutation databases at http://www.hgvs.org/content/databases-tools. For further background see PMID 17893115.

7.3 PATHOGENESIS DUE TO VARIATION IN SHORT TANDEM REPEAT COPY NUMBER

Our genome is littered with short tandem repeats of mononucleotides to oligonucleotides. Very few, notably TTTAGG repeats in telomere DNA, are functionally important. The great majority are there simply as a result of statistical inevitability. In our genome, which has a 41 % GC/59 % AT base composition, the odds that a hexanucleotide chosen at random has the sequence AAAAAA is 1 in $(0.295)^6$ or 1 in 1,517. In an average human chromosome (with 133 Mb of DNA) the sequence AAAAAA occurs over 87 000 times; nearly 16 000 examples of the decanucleotide ATATATATAT would be expected across the whole genome.

Short tandem repeats cause problems for aligning the two opposing DNA strands of a double helix. They increase the chances of *slipped strand mispairing*, the local mispairing of repeats on opposing DNA strands, causing the strands to slip slightly out of alignment. For an ATATATATAT array, for example, the third AT on one strand might mistakenly base pair with the fourth AT on the opposing strand, with consequent looping out of an AT. If this happens during DNA replication, then the newly synthesized DNA strand will have fewer or more AT repeats. We previously covered the general principle of **replication slippage** when we explained the basis of simple tandem repeat micropolymorphism (Figure 4.6 on page 91). DNA strands may also be produced with fewer or more tandem repeats when DNA repair occurs at a DNA region where there is slipped strand mispairing.

The two main classes of pathogenic variation in short tandem repeat copy-number

Length variation in arrays of short tandem repeats results in disease in two ways, as listed below.

- *Frameshifting expansion/contraction in coding DNA*. Replication slippage can cause very short frameshifting insertions or deletions within arrays of short tandem repeats in coding DNA (the great majority are 1-or 2-nucleotide insertions/deletions). They account for a very significant component of pathogenic insertions and deletions in coding DNA.
- *Non-frameshifting expansion beyond safe limits*. Disease can occur when the number of repeats in some short tandem repeats expands beyond some safe limit, causing moderately long to quite large insertions at the DNA level. The expansions occur for certain triplet repeats in coding DNA and various types of repeat in noncoding DNA,

As an example of how very short pathogenic insertions and deletions are produced through replication slippage in short tandem repeats, consider the tandem mononucleotide sequence AAAAAA occurring on the sense strand of a coding DNA, and replication slippage leading to a single A being lost or gained (to give five or seven adenines, respectively). The loss (deletion) or gain (insertion) of an adenine would produce a frameshift in the translational reading frame, often resulting in a premature termination codon and failure to make a protein. Loss or gain of repeats within arrays of tandem di-, tetra- and pentanucleotides in coding DNA can similarly cause disease by introducing frameshifting and early premature termination codons.

Trinucleotide repeats in coding DNA are different: loss or gain of a single repeat does not affect the translational reading frame, and often the effect of the changed length is inconsequential. But non-frameshifting expansion of certain triplet repeats can nevertheless be pathogenic, sometimes by inactivating a gene, or by causing proteins to behave abnormally or by producing toxic RNAs, as described below.

Non-frameshifting pathogenic expansion in short tandem repeat number

More than 40 diseases are caused by expansion of the number of short tandem repeats (from trinucleotide to dodecanucleotide repeats) beyond safe limits. Some occur in coding DNA. Others are located in 5¢ or 3¢ untranslated sequences or in introns within a protein-coding gene. In the former case certain types of tandem triplet repeats are involved that ultimately specify alanine or glutamine (polyalanine and polyglutamine tracts are found in a considerable number of human proteins). Pathogenic short tandem repeat expansion in noncoding DNA involves different types of tandem repeats, from triplet to dodecanucleotide repeats. In the sections that follow we detail some specific examples. The list below provides an overview.

- *Polyalanine expansion*. Nine congenital disorders of development are known to be caused by expansion of alanine-specifying triplet repeats. (The resulting polyalanine tracts typically occur in transcription factors, serving as flexible nonpolar linkers between two folded domains.) Individual arrays often contain different types of alanine-specifying triplets, and are not polymorphic, while quite small, stable expansions are seen in affected individuals (**Table 7.7**). Disease results after the expansions pass a certain safe-size limit: proteins with a sufficiently extended polyalanine tract can

aggregate to cause problems for cells. We examine protein aggregation more generally in Section 7.8.

TABLE 7.7 EXAMPLES OF DISEASES RESULTING FROM THE THREE CLASSES OF NON-FRAMESHIFTING PATHOGENIC EXPANSION OF SHORT TANDEM REPEATS

Class (Stability*)	Disease examples	Repeat unit	Associated tandem repeats		
				Copy number	
			Location	Normal	Disease
Polyalanine expansion (STABLE)	Oculopharyngeal muscular dystrophy (OPMD)	GCG, GCT, GCA (specifying alanine)	Coding sequence	10	11-17
	Synpoldactyly type II			15	22-29
	Hand-foot genital syndrome			18	24-26
Polyglutamine expansion (UNSTABLE)	Huntington disease	CAG (specifying glutamine)	Coding sequence	6-35	36-250
	Spinal bulbar muscular atrophy (Kennedy disease)			4-34	35-72
	Dentatorubropallidoluysian atrophy			3-38	49-88
	Spinocerebellar ataxia type 7			7-41	43-51
Noncoding DNA expansion (UNSTABLE)	Friedreich ataxia	GAA	Intronic	6-32	200-1700
	Fragile X syndrome (with intellectual disability)	CGG	5’ - UTR	5-54	>200 to several ×1000
	Myotonic dystrophy type I	CTG	3’ - UTR	5-37	50-10000
	Myotonic dystrophy type II	CCTG	3’ - UTR	10-26	75-11 000

* Some short tandem repeat expansions are unstable (see text).

Class (Stability*)	Disease examples	Repeat unit	Associated tandem repeats		
				Copy number	
			Location	Normal	Disease
	Spinocerebellar ataxia type 10	ATTCT	Intronic	10-29	500-4500
	Fontal dementia and/or amyotrophic lateral sclerosis	GGGGCC	Intronic	2-22	700-1600

* Some short tandem repeat expansions are unstable (see text).

- *Polyglutamine expansion*. Nineteen disorders with neurodegenerative or neuromuscular phenotypes show expansion of CAG repeats specifying polyglutamine. Polyglutamine tracts are highly flexible, but unlike polyalanine are highly polar. Unlike the triplet repeats associated with polyalanine expansion, those associated with polyglutamine expansion are often homogenous CAG repeats showing length polymorphism in the general population but moderate to sometimes quite large expansions in affected individuals (see Table 7.7). The pathogenic expansion in polyglutamine tracts can be quite unstable, increasing in size after both mitotic and meiotic cell division (the underlying mutations are described as *dynamic mutations*—see next section).
- *Pathogenic expansion of noncoding short tandem repeats*. More than 20 human disorders are due to moderate to large expansion of short tandem repeats in introns or untranslated sequences of protein-coding genes. The repeat units are mostly 3 to 6 nucleotides in length—see Table 7.7 for examples. These expansions can show very significant instability in mitotic and meiotic cell division (see next section).

Dynamic disease-causing mutations due to unstable expansion of short tandem repeats

We are accustomed to thinking that mutations are stable. When a disease-causing mutation is transmitted from one generation to the next, we expect to see the same mutation in affected individuals from different generations of a family. However, pathogenic expansion of polyglutamine-specifying CAG repeats and various types of

short tandem repeats in noncoding DNA can be unstable. They are sometimes described as **dynamic mutations** because the repeat length can increase from one generation to the next (and sometimes from mother cell to daughter cell in one individual). In marked contrast, in polyalanine expansion disorders the triplet repeat expansions are stable.

Increasing expansion of the short tandem repeats leads to increasing severity of the disease. Because the expansions can increase from one generation to the next, the phenotype can become increasingly severe from one generation to the next, a phenomenon known as **anticipation**. We previously gave an example of anticipation in myotonic dystrophy wherein the disease progressed from mild features in a grandmother to moderate features in her daughter and then to severe congenital muscular dystrophy in the grandson (see Figure 5.17 on page 129). And because disorders arising from unstable expansion of short tandem repeats are often neurodegenerative, pronounced differences in the age of onset of symptoms can be attributable to the extent of repeat expansion, as shown strikingly in the case of Huntington disease (**Figure 7.6**).

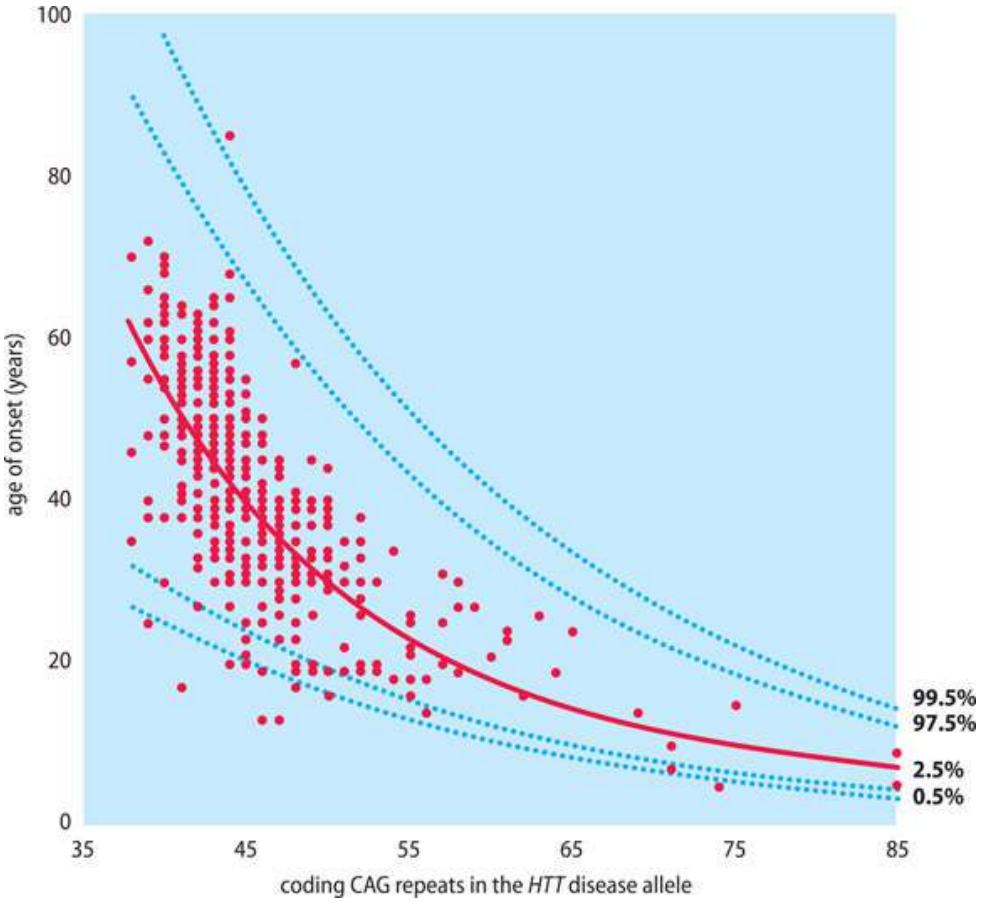


Figure 7.6. Inverse correlation between the extent of CAG repeat expansion and age of disease onset in Huntington disease. An increased CAG repeat number means increased loss of neurons and earlier evidence of symptoms. Reproduced from Budworth H, McMurray CT (2013) A brief history of triplet diseases. In: Kohwi Y, McMurray C (Eds) *Trinucleotide repeat protocols. Methods in molecular biology (methods and protocols), 1010*. Humana Press, Totowa, NJ. With permission from Springer Nature.

The unstable and large expansion of polyglutamine-specifying CAG repeats and certain noncoding tandem repeats is in marked contrast to the limited and very stable expansion of polyalanine-specifying triplet repeats. The difference is likely to be due to the comparative heterogeneity of polyalanine-specifying triplet repeats (individual arrays are heterogeneous, having two or more different triplets in arrays, hindering slipped strand mispairing and so limiting expansion). Arrays of polyglutamine-specifying CAG repeats and noncoding tandem repeats generally show very high repeat homogeneity.

The mechanism underlying unstable tandem repeat expansion is unclear. The arrays of tandem repeats are prone to forming abnormal secondary structures, including slipped strands with extrahelical loops, and hairpins, and also certain other unusual structures, including triplex and quadruplex DNA. Secondary structure elements such as these can impede DNA functions (replication, transcription, and so on), provoking DNA repair mechanisms to remove them. From studies of both mouse models and human genomewide association, components of the mismatch DNA repair system are known to be involved in the unstable repeat expansions.

The size of unstable repeat expansions can be very large for some disorders, notably for myotonic dystrophy (types 1 and 2) and Fragile X. As a result, for these disorders, a specialized laboratory analysis, known as triplet-repeat primed PCR is used to track the repeat size (see **Clinical Box 5** for a case study of myotonic dystrophy, showing an example of this analysis).

The complication of repeat-associated non-AUG translation

At the DNA level, unstable tandem repeat expansions might seem broadly similar, but at the cell level the pathogenic consequences can differ remarkably, as outlined in the subsection following this one. There is the added complication of unorthodox repeat-specific expression mechanisms as well as orthodox expression. Expanded CAG repeats can be conventionally expressed to give large polyglutamine tracts, but

are also subject to unorthodox *repeat-associated non-AUGtranslation* (*RAN translation*) in multiple reading frames to also produce polyserine (specified by recurring AGC) and polyalanine (specified by recurring GCA).

RAN translation across transcripts from expanded tandem DNA repeats does not require an AUG start codon and applies to expanded tandem noncoding DNA. Both sense and antisense transcripts from expanded noncoding repeats can be RAN-translated in different reading frames (see **Figure 7.7**). The contribution of RAN-translation expression products to pathogenesis is an area of active investigation.

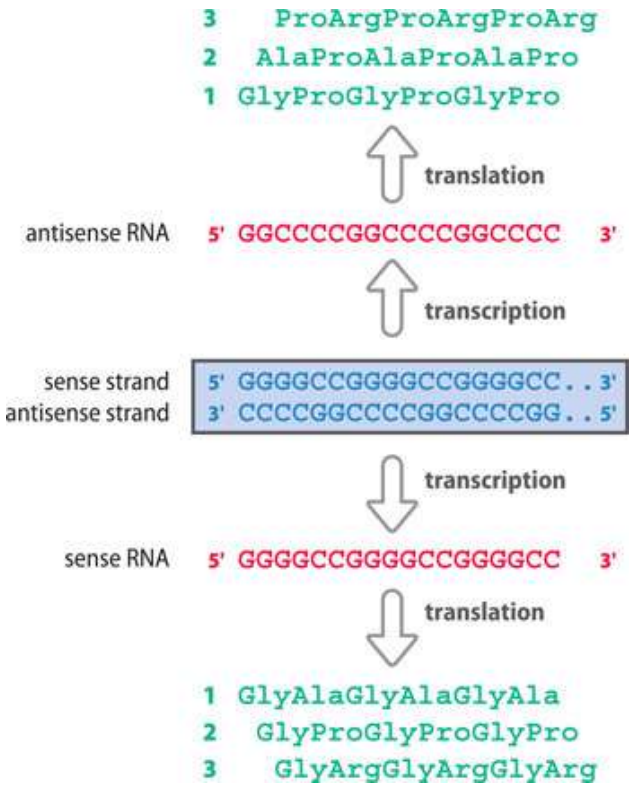


Figure 7.7.RAN translation produces five different polypeptides from tandem GGGGCC repeats associated with frontal dementia and amyotrophic lateral sclerosis. Transcription of both strands of tandem GGGGCC repeats (also called G_4C_2 repeats) produces both sense and antisense RNA transcripts for each of which translation can occur in all three reading frames to give five different polypeptide products: poly (Gly-Ala); poly (Gly-Pro); poly (Gly-Arg); poly (Pro-Arg); and poly (Ala-Pro).

CLINICAL BOX 5 A CLINICAL CASE STUDY: MYOTONIC DYSTROPHY TYPE I

Janna was born at term from an uneventful pregnancy to non-consanguineous parents. She had a normal development through childhood. In early adolescence she had difficulty in relaxing her hand muscles after making a fist, and in opening her mouth after the first bite of a meal. Janna also had difficulty in keeping focused on one task for more than an hour. She felt continuously tired, and would regularly sleep 12 hours a night or longer when not using an alarm. Soon after, she started having bouts of diarrhea that would last for some days, followed by episodes of constipation and abdominal pain. As an adult, Janna developed difficulties walking; she tripped constantly when walking on uneven surfaces outdoors and experienced several falls per month. Her voice changed, becoming nasal. She increasingly had difficulties swallowing, and eventually experienced some choking episodes. She was initially referred to an orthopedic surgeon who requested an EMG that identified myotonic discharges in several muscles (Figure 1). She was then referred to the neuro-muscular center for further follow-up.

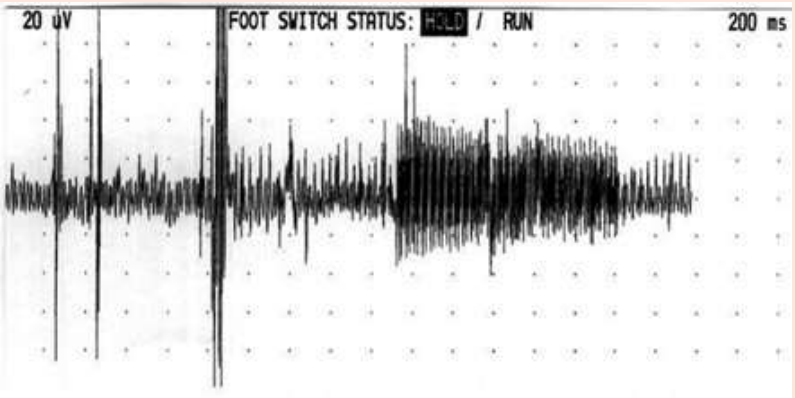


Figure 1 Electromyography (EMG) showing myotonic discharges. Several of Janna's muscles showed spontaneous discharges of muscle fibers that waxed and waned in both amplitude and frequency.

First seen by the neuromuscular team at the age of 25, Janna presented with mild ptosis, facial weakness and atrophy of temporal and sternocleidomastoid muscles (Figure 2). She had mild weakness of finger flexor muscles, myotonia in the hands, and distal muscle weakness in her lower limbs, affecting both tibialis anterior (3-/5 on the MRC scale).

Figure 2 The typical myopathic face in myotonic dystrophy. Janna had an elongated typical myopathic face, temporal muscle atrophy, ptosis, perioral weakness, a high arched palate and bilateral atrophy of the sternocleidomastoid muscles.

Janna was diagnosed with myotonic dystrophy type 1 in the genetic lab after TP-PCR analysis showed she had an abnormal expansion of >350 CTG repeats in the *DMPK* gene (**Figure 3**). Follow-up testing of Janna's parents and siblings confirmed a genetic diagnosis in her father, John (who had only minor symptoms, and did not show myotonia or muscle weakness). Janna's elder brother, Mark, was found to have myotonia in his hands and jaw, but no muscle weakness, but her younger brother, George, was unaffected.

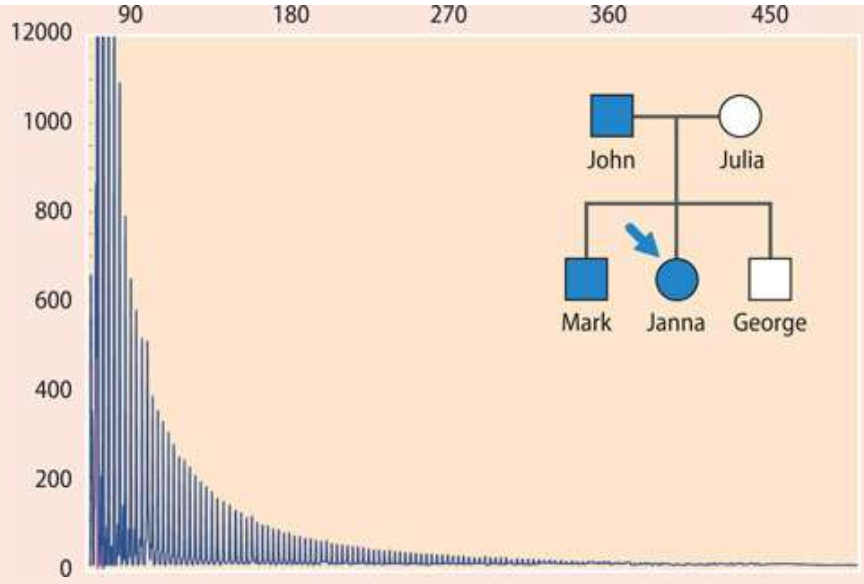


Figure 3 Triplet repeat-primed PCR (TP-PCR) result. It shows Janna has >350 CTG repeats in exon 15 of the *DMPK* gene (compared to 50 repeats or less in unaffected controls). The inset shows the pedigree obtained after following up Janna's family.

Janna underwent counselling and was informed that there was a 50 % risk of passing her condition to her off-spring. She was advised that due to the phenomenon of anticipation, an affected child would be at high risk of presenting with the most severe form of myotonic dystrophy, congenital myotonic dystrophy. A baby affected in this way would be expected to be floppy and might experience severe breathing and feeding difficulties after birth. And the instability of the repeat expansion would be expected to increase during life, and be most pronounced in post-mitotic cells, such as in muscle, brain, and heart.

Unstable expansion of short tandem repeats can cause disease in different ways

Unstable short tandem repeat expansion can cause loss of function in recessive disorders (carrier individuals with one normal allele are unaffected). In other cases, a mutant gene bearing the expanded repeat is expressed and the problem is a gain of function: the gene makes an aberrant RNA or protein product that is harmful to cells and causes disease in heterozygotes (the presence of a normal allele is insufficient to prevent the disease phenotype)—see **Table 7.8**.

TABLE 7.8 THREE MAJOR PATHOGENESIS ROUTES FOR UNSTABLE TANDEM REPEAT

EXPANSION DISORDERS			
Mutation class	**Cause of disease**	**Example of disorder**	**Comments**
Loss of function	Gene expression is inhibited; not enough gene product is made	Fragile X syndrome	Expansion of a tandem CGG repeat in exon 1 of the *FMR1* geneto > 200 repeats triggers methylation of the promoter region and prevents transcription.
Gain of function	Harmful RNA transcript	Myotonic dystrophy (types 1 and 2)	Abnormally expanded CUG or CCUG repeats in the 3'- UTR of mRNA sequester regulatory proteins that control alternative splicing (see text).
	Cytotoxic protein	Huntington disease	The CAG expansion causes production of a huntingtin protein with an abnormally large number of glutamines that is especially toxic to neurons

For both myotonic dystrophy types 1 and 2, the problem is that RNA transcripts containing expanded tandem repeats bind to, and sequester, MBNL1 and other members of the muscleblind-like family of regulatory proteins that control alternative RNA splicing. These proteins have binding sites for both CUG and CCUG, and so expanded CUG repeats and expanded CCUG repeats in transcripts from the disease gene (in patients with myotonic dystrophy type 1 and type 2 respectively), bind to large quantities of the MBNL family of regulators, preventing their normal function. As a result, cells have a plethora of splicing defects at various gene loci.

The third pathogenic class, characterized by production of a toxic protein, is known to be common in the case of unstable CAG repeat expansion in coding sequences. Proteins with large polyglutamine tracts are unstable, and prone to aggregation in a way that is toxic to cells. Because neurons are intended to be extremely long-lived cells they are not readily replaced, and steady neuron depletion over time results in neurodegenerative and neuromuscular disorders.

Disease manifestation in Fragile X pre-mutation carriers

The CGG repeat expansion in exon of the *FMR1* gene is unusual because it results in different diseases according to the extent of the expansion. The "full mutation" that causes Fragile X syndrome is defined as expansions of over 200 CGG repeats. When this 200-repeat limit is passed, methylation of the nearby promoter is triggered, and the *FMR1* gene, which is important in brain function, is silenced, causing cognitive defects in males. (Note: because it is a loss-of-function phenotype, a small percentage of cases may occur by alternative point mutations and deletions in *FMR1*.)

Abnormal expansion of the CGG repeat array to lesser levels of 55–200 repeats (*pre-mutations*) also causes disease but by a different, poorly misunderstood gain-of-function mechanism in which *FMR1* mRNA is produced in excess, causing toxicity and mitochondrial dysfunction. 40 % of male carriers of this pre-mutation and 16 % of female carriers develop Fragile X-associated tremor/ataxia syndrome (FXTAS), with late onset (typically between 60 and 65 years of age) of progressive cerebellar ataxia and intention tremor, followed by cognitive impairment. In addition, 20 % of women with the premutation allele develop Fragile X-associated primary ovarian insufficiency (FXPOI) with hypergonadotrophic hypogonadism before the age of 40 years (compared to 1 % in the general population).

7.4 PATHOGENESIS TRIGGERED BY LONG TANDEM REPEATS AND INTERSPERSED REPEATS

Disease caused by various larger changes in DNA may involve large scale changes, facilitated by long tandem repeats and interspersed repeats. Inappropriate pairing up of non-allelic but closely homologous repeats in nuclear DNA and subsequent recombination can produce pathogenic duplications, deletions and inversions, and this is the area that most of this section is devoted to. (Of course, very large deletions, duplications and inversions, many of which may occur by the same mechanisms, have long been recorded under the microscope; we cover these chromosomal abnormalities together with chromosome segregation abnormalities in Section 7.5.).

Pathogenic exchanges between repeats occurs in both nuclear DNA and mtDNA

Many of the repeats involved in triggering large-scale mutations occur within introns or outside genes. Some large repeats, however, contain one or more gene sequences. Pathogenesis may arise when large-scale mechanisms adversely change the structure or copy number of genes, or when they adversely alter gene expression. We detail the principal genetic mechanisms giving rise to these changes in individual sections below.

The non-allelic but closely homologous repeat sequences that predispose to moderate to large-scale deletions, insertions and inversions in nuclear DNA belong to two classes as listed below.

- *Long tandem DNA repeats.* Recall from Section 2.5 that local DNA duplication events have frequently occurred during genome evolution. The most recent produced highly homologous tandem repeats many kilobases or megabases long and often containing multiple genes (*segmental duplication*). The classic example of involvement in pathology is provided by a tandem repeat of ~30 kb that includes the steroid 21-hydroxylase gene: pairing of nonallelic repeats at meiosis is responsible for 99 % of the pathology in 21-hydroxylase deficiency, as detailed below.
- *Interspersed repeats.* As well as the very high copy-number interspersed Alu and LINE repeats, many families of highly homologous low copy-number repeats are found in noncoding DNA in and around genes. Large genes with many long introns have multiple interspersed repeats in the introns, making them more susceptible to non-allelic mispairing at meiosis. The dystrophin gene has 78 introns of average size >30 kb, and so it is not surprising that ~75 % of boys with Duchenne muscular dystrophy have large pathogenic deletions or duplications, many causing a shift in the translational reading frame (where the *net* effect is to delete or insert a number of coding nucleotides that is not a multiple of three; Figure 2C in Box 2.1 on page 27 shows the general concept). Moderate insertions can also occur within a coding sequence to produce an extended protein that might be unstable, not fold properly, or be functionally disadvantaged). As detailed in Section 7.6, deletions occur quite frequently in mtDNA, very often arising by sequence exchanges between interspersed repeats.

Non-allelic homologous recombination and transposition

Large-scale pathogenic mutations in nuclear DNA are often initiated by abnormal pairing between low-copy-number repeats with very similar sequences (*homologous repeats*), often occurring in noncoding DNA within genes, or close to the genes. Different families of low-copy-number repeats can be distinguished. They might be short interspersed sequences, or naturally duplicated sequences extending up to several hundreds of kilobases in length and containing multiple genes.

When two repeats with very similar sequences occur close to each other on the same chromosome arm, the high level of sequence identity between the repeats can lead to mispairing of chromatids. Non-allelic repeats can then pair up: repeat no. 1 on one chromatid might pair up with repeat no. 2 on the other chromatid. A subsequent recombination event occurring in the mispaired region produces a change in repeat copy number; the process is known as **nonallelic homologous recombination** (**NAHR**).

Alternatively, recombination occurs between homologous repeats on the same DNA molecule; an *intrachromatid recombination* such as this is also a form of NAHR. We describe different NAHR mechanisms below, show how they can generate insertions, deletions, and inversions, and give examples of how they cause disease.

Pathogenic sequence exchanges between chromatids at mispaired tandem repeats

Many human genes and gene regions have significant arrays of long tandemly repeated DNA sequences. This can include the repetition of exons, whole genes, and even multiple genes. Tandem repeats within genes or spanning coding sequences can predispose to a type of nonallelic homologous recombination that can cause disease.

Normally, recombination between homologous chromosomes occurs after the chromosomes have paired up with their DNA sequences in perfect alignment. However, local misalignment of the paired chromosomes is more likely to occur in regions where there are highly similar tandem repeats—the DNA molecules of the two chromatids can line up out of register. That is, the alignment is staggered and one or more repeats on each chromatid do not pair up with their normal partner repeat on the other chromatid.

A subsequent recombination within the mismatched sequences is known as **unequal crossover** (**UEC**) and results in one chromatid with an insertion (more tandem repeats) and one with a deletion (fewer tandem repeats) An equivalent process can also occur between sister chromatids, an *unequal sister chromatid*

exchange (UESCE) (Figure 7.8). UEC and UESCE cause reciprocal exchanges between misaligned chromatids: one chromatid gains an extra DNA sequence, and the other loses an equivalent sequence. Disease may result from a change in gene copy number, or through the formation of hybrid genes that lack some of the functional gene sequence.

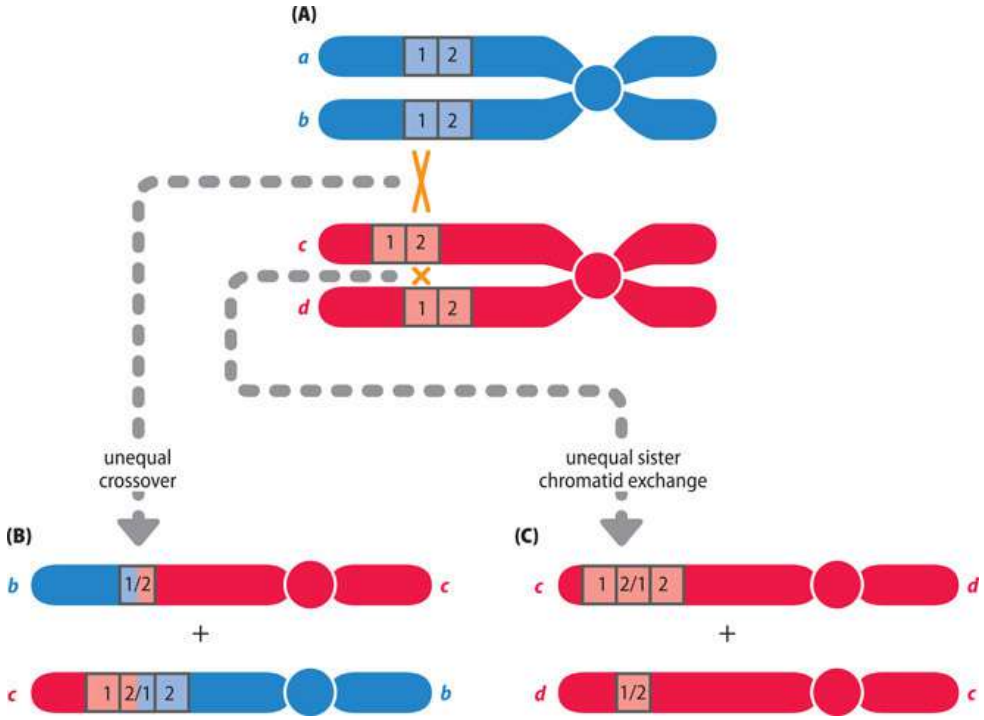


Figure 7.8 Unequal crossover and unequal sister chromatid exchange cause deletions and insertions. (A) Mispairing of chromatids with two very similar tandem repeats (1, 2). The very high sequence identity between these repeats can facilitate misalignment between the DNA of aligned chromatids so that repeat 1 on one chromatid aligns with repeat 2 on another chromatid. The misaligned chromatids can be on non-sister chromatids of homologous chromosomes, such as chromatids *b* and c, in which case recombination (large orange X) in the misaligned region results in an unequal crossover (B). Alternatively, there can be a recombinationlike unequal sister chromatid exchange (small orange X) between misaligned repeats on sister chromatids (c, d) of a single chromosome (C). In either case, the result is two chromatids, one with three repeat units and the other with a single repeat unit (a hybrid of sequences 1 and 2—shown here as 1/2 or 2/1).

Misalignment of repeats on paired chromatids can also cause disease by *non-reciprocal* sequence exchange. Here, one of the interacting sequences remains unchanged, but the other is mutated (**gene conversion**—see Figure 7.9). See **Clinical Box 6** for a common single-gene disorder, steroid 21-hydroxylase

deficiency, in which the pathogenesis is due almost entirely to sequence exchanges between misaligned long tandem repeats, resulting in deletion or gene conversion.

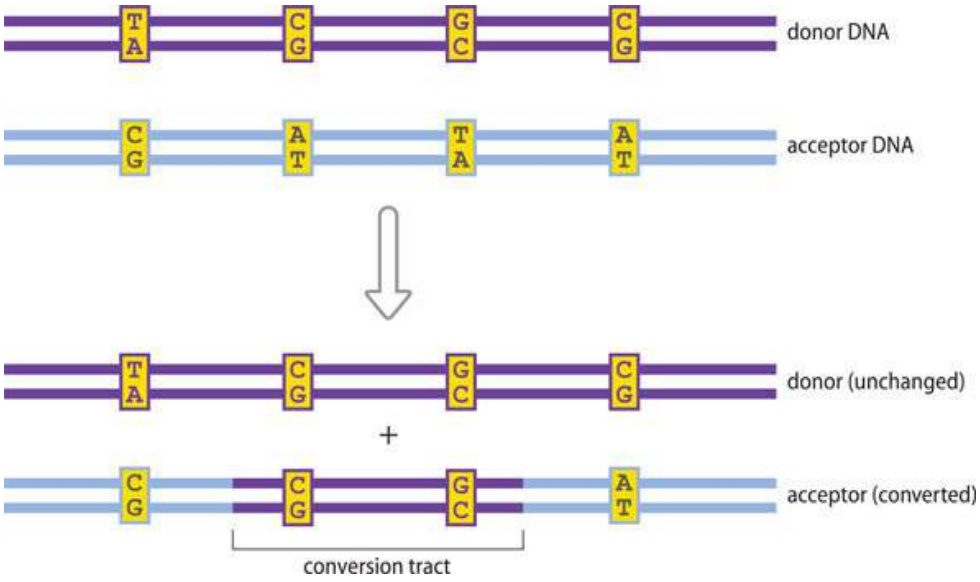


Figure 7.9 Principle of gene conversion. Gene conversion is a *nonreciprocal* sequence exchange between two related sequences that may be alleles or non-allelic (such as misaligned repeats on opposing non-sister chromatids). Sequence information is copied from one of the paired sequences, the donor sequence (which will remain unchanged) to replace an equivalent part of the other sequence, the acceptor sequence. The size of the sequence that is converted—the conversion tract—is often a few hundred nucleotides long in mammalian cells.

CLINICAL BOX 6 DISEASE PROFILE: STEROID 21-HYDROXYLASE DEFICIENCY, A DISORDER CAUSED BY GENE-PSEUDOGENE SEQUENCE EXCHANGES

Steroid 21-hydroxylase is a cytochrome P450 enzyme needed by the adrenal gland to produce the glucocorticoid hormone cortisol and aldosterone (which regulates sodium and potassium levels). Genetic deficiency in this enzyme is much the most common cause of congenital adrenal hyperplasia. In classical (congenital) forms of the disorder, excessive adrenal androgen bio-synthesis results in virilization of affected individuals, so that girls are often born with masculinized external genitalia. Classically affected individuals may have the “simple-virilizing” form of the disorder, but some also excrete large amounts of sodium in their urine, which leads to potentially fatal electrolyte and water imbalance (“salt-wasting” phenotype).

Steroid 21-hydroxylase is encoded by a gene, *CYP21A2,* located in the class III region of the HLA complex. *CYP21A2* resides on an approximately 30 kb segment of DNA that is tandemly duplicated, with about 98 % sequence identity between the tandem 30 kb repeats. As a result, there is a closely related copy of the *CYP21A2* gene sequence on the other repeat, a pseudogene called *CYP21A1P*. The pseudogene has multiple deleterious mutations distributed across its length (**Figure 1A**) and does not make a protein.

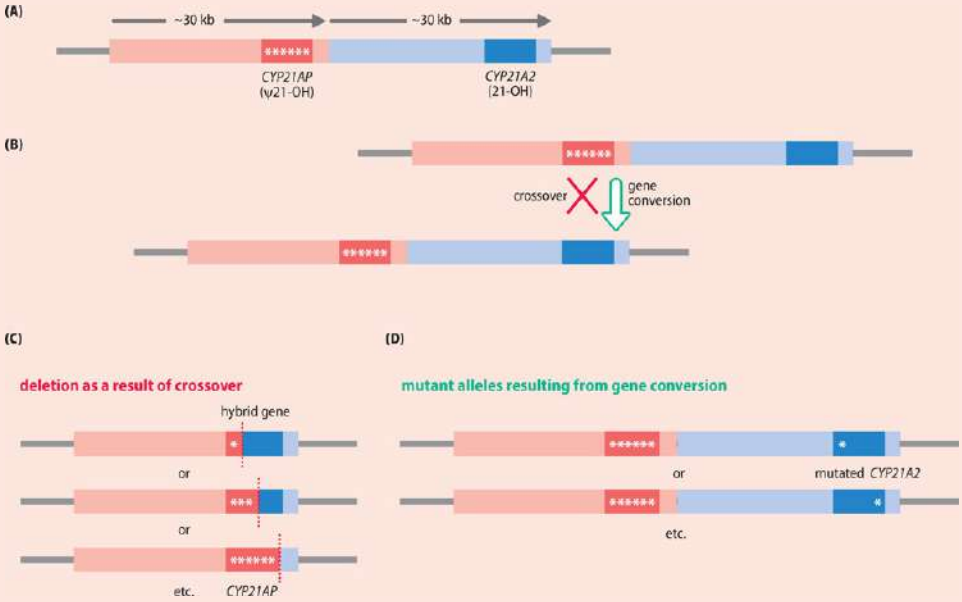


Figure 1 Tandem duplication of repeats containing the *CYP21A2* gene and a closely related pseudogene predispose to pathogenic sequence exchanges. (A) The steroid 21-hydroxylase gene, *CYP21A2,* and a closely related pseudogene, *CYP21AP* (with multiple inactivating mutations shown by white asterisks), are located on tandemly duplicated ~30 kb repeats that also contain other genes (not shown here for simplification). (B) Mispairing of the non-allelic repeats is facilitated by the 98 % sequence identity between them. A subsequent reciprocal exchange occurring by recombination (shown by the red X) or a nonreciprocal exchange (gene conversion) can result in a loss of functional sequence. (C) Unequal crossover can produce chromosomes with a single 30 kb repeat containing a hybrid gene, part *CYP21AP* and part *CYP21A2,* or just the *CYP21AP* pseudogene (dashed vertical red lines indicate crossover points). (D) Gene conversion replaces a segment of the *CYP21A2* gene by equivalent sequence from the pseudogene. The conversion tract, often only a few hundred nucleotides long, can occur at different regions, introducing a copy of an inactivating mutation from any part of the pseudogene.

The very high sequence identity between the tandem repeats makes paired chromatids liable to local misalignment: a repeat containing the functional

CYP21A2 gene on one chromatid mispairs with the repeat containing the *CYP21AP* pseudogene on the other chromatid (Figure 1B).

In more than 99 % of cases with 21-hydroxylase deficiency, mispairing between the two repeats and subsequent sequence exchanges is thought to be responsible for pathogenesis. About 75 % of the mutations that cause disease are roughly 30 kb deletions caused by unequal crossover or unequal sister chromatid exchange. If the crossover point occurs between gene and pseudogene, a single nonfunctional hybrid 21-hydroxylase gene results; or if it is located just beyond the paired gene and pseudogene, it leaves just the 21-hydroxylase pseudogene (Figure 1C).

The remaining 25 % or so of pathogenic mutations are point mutations, but in the vast majority of cases the point mutation is introduced into the *CYP21A2* gene by a gene conversion event that copies a sequence containing a deleterious mutation from the pseudo-gene, replacing the original sequence (Figure 1D and **Table 1**).

TABLE 1 PATHOGENIC POINT MUTATIONS IN THE STEROID 21-HYDROXYLASE GENE ARE COPIED FROM A CLOSELY RELATED PSEUDOGENE

Mutation class and location	**Normal 21 -OH gene sequence *(CYP21A2)***	**Pathogenic point mutation**	**Equivalent *CYP21A2P* pseudogene sequence**
Intron 2, splicing mutation	CCCACCCTCC	CCCA**G**CCTCC	CCCA**G**CCTCC
Exon 3, deletion of 8 ntds. within codons 111-113	GGA GAC TAC TCx Gly Asp Tyr Ser	G-- --- --- TCx **V()al**	G-- --- --- TCx
Exon 4, missense mutation I173N	ATC ATC TGT Ile Ile Cys	ATC A**A**C TGT Ile Asn Cys	ATC A**A**C TGT
Exon 4, clustered missense mutations in codons237-240	ATC GTG GAG ATG Ile Val Glu Met	A**A**C G**A**G GAG A**A**G **Asn Glu** Glu **Lys**	A**A**C G**A**G GAG A**A**G
Exon 7, missense mutation:V282L	CAC GTG CAC His Val His	CAC **T**TG CAC His Leu His	CAC **T**TG CAC

The gene conversion tract (the region copied from the pseudogene sequence) is usually no longer than a few hundred nucleotides.

Mutation class and location	Normal 21 -OH gene sequence *(CYP21A2)*	Pathogenic point mutation	Equivalent *CYP21A2P* pseudogene sequence
Exon 8, nonsense mutation: Q319X	CTG CAG GAG Leu Gln Glu	CTG TAG GAG Leu **STOP**	CTG **T**AG GAG
Exon 8, missense mutation: R357W	CTG CGG CCC Leu Arg Pro	CTG **T**GG CCC Leu **Trp** Pro	CTG **T**GG CCC

The gene conversion tract (the region copied from the pseudogene sequence) is usually no longer than a few hundred nucleotides.

Disease arising from sequence exchanges between distantly located repeats in nuclear DNA

Homologous repeats that are separated by a sizable intervening sequence can also predispose to nonallelic homologous recombination in which sequence exchanges occur between misaligned repeats. The repeats may be **direct repeats** (oriented in the same 5¢ ® 3¢ orientation); in that case, the intervening sequence between the repeats can be deleted or duplicated. That may mean loss or duplication of exons, which can be frameshifting mutations, or loss or duplication of multiple genes (both of which can be pathogenic, as described below). Exchange between *inverted repeats* (repeats oriented in opposite 5¢ ® 3¢ directions) can also cause inversion of the intervening sequence; this can also be pathogenic, as illustrated in the last subsection below.

Chromosome microdeletions and microduplications

Just as with mispairing of tandem repeats, distantly spaced direct repeats can mispair when chromatids are aligned within homologous chromosomes. Subsequent recombination at mismatched direct repeats results in deletions or duplications of the intervening sequence (**Figure 7.10A**). An equivalent type of exchange can also occur between mispaired short direct repeats on the *same* DNA strand (a form of intrachromatid recombination), and this can also lead to deletion (Figure 7.10B).

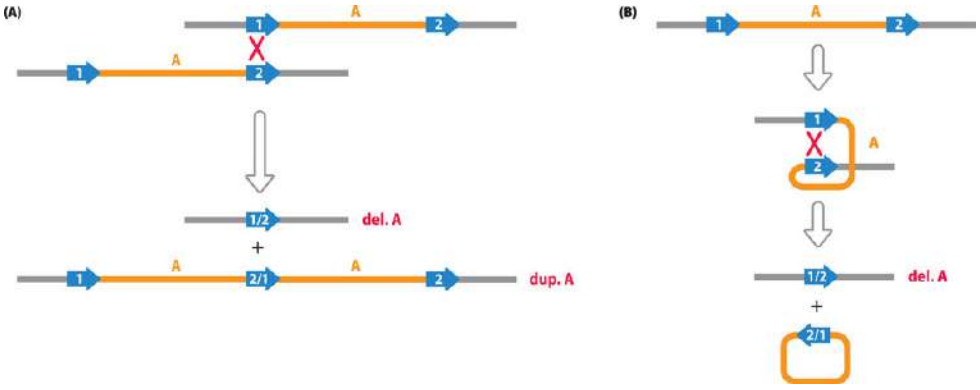


Figure 7.10 Deletion/duplication events due to nonallelic homologous recombination between low-copy-number direct repeats. Block arrows indicate highly similar low-copy-number *direct repeats* (oriented in the same direction). (A) DNA sequence A is flanked by low-copy-number repeats 1 and 2 that have identical or highly similar nucleotide sequences. Chromatid misalignment can occur so that repeat 1 on one chromatid pairs with repeat 2 on the other chromatid, and subsequent crossover can result in a chromatid with a deletion of sequence A (del. A) and one with two copies of sequence A (dup. A). (B) An intrachromatid "recombination" between direct repeats on the same DNA molecule can also produce a deletion of sequence A. If the other product, a circular DNA containing sequence A, lacks a centromere, it will be lost after cell division.

The sizes of the duplication/deletions produced are frequently <6 Mb (see **Table 7.9**), and because they are not detectable under the light microscope using standard chromosome staining they have been described as chromosome *microdeletions* or *microduplications*. As shown in Table 7.9 several examples are known to be associated with clinical phenotypes. We consider the effects on gene expression and the clinical impact in Section 7.5.

TABLE 7.9 EXAMPLES OF CLINICAL PHENOTYPES ARISING THROUGH RECOMBINATION BETWEEN DISPERSED REPEATS ON A CHROMOSOME

Disorder	Location	Length of recombining repeats	Deletion (Δ) / duplication (dup.) size	Key disease locus
Azoospermia type AZFc	Yq11.2	230 kb	Δ 3.5 Mb	*DAZ family*
Angelman syndrome	15q11–q13	400 kb	paternal Δ 5 Mb	*UBE3A*

Abbreviations are: HNPP—hereditary neuropathy with liability to pressure palsies; VCFS—velocardiofacial syndrome.

Disorder	Location	Length of recombining repeats	Deletion (Δ) / duplication (dup.) size	Key disease locus
Prader-Willi syndrome			maternal Δ 5 Mb	*SNRPN*
HNPP	17p12	24 kb	Δ 1.4 Mb	PMP22
Charcot-Marie-Tooth 1A			dup. 1.4 Mb	
DiGeorge syndrome/VCFS	22q11.2	225–400 kb	Δ 3 Mb or 1.5 Mb	*TBX1*
Smith-Magenis syndrome	17p11.2	175–250 kb	Δ 4 Mb	RAI1
Potocki-Lupski syndrome			dup. 4Mb	
Williams-Beuren syndrome	7q11.2	300–400 kb	Δ 1.6 Mb	*ELN*
Sotos syndrome	5q25	400 kb	Δ 2 Mb	NSD1

Abbreviations are: HNPP—hereditary neuropathy with liability to pressure palsies; VCFS—velocardiofacial syndrome.

Intrachromatid recombination between inverted repeats

Inverted repeats on a single chromatid can also mispair by looping out the intervening sequence. Subsequent recombination at the mispaired sequences will produce an inversion of the intervening sequence (**Figure 7.11A**). Disease may result, for example by relocating part of a gene, by disrupting the gene, or by separating a gene from important *cis-acting* control elements.

An instructive example is provided by hemophilia A: in about 50 % of cases the cause is a large inversion that disrupts the *F8* gene, which makes blood clotting factor VIII. A low-copy-number repeat, *F8A1,* within intron 22 of the *F8* gene can mispair with either of two very similar repeat sequences, *F8A1* and *F8A2,* that are located upstream of *F8* and in the opposite 5¢ ® 3¢ orientation to *F8A1*. Subsequent

recombination between mispaired *F8A* repeats produces an inversion of about 500 kb of intervening sequence, causing disruption of the *F8* gene (Figure 7.11B).

7.5 CHROMOSOME ABNORMALITIES

Many large-scale changes to our DNA sequences that cause diseases are more readily studied at the level of chromosomes, as are changes in chromosome copy number resulting from errors in chromosome segregation. In standard cytogenetic **karyotyping** suitable metaphase or prometaphase chromosome preparations are chemically stained to reveal a pattern of alternating light and dark bands under light microscopy, which are examined to identify chromo-some abnormalities. See **Box 7.2** for some details of the techniques and relevant nomenclature. Alternative approaches to identify or screen for chromosome abnormalities are detailed in Chapter 11.

BOX 7.2 HUMAN CHROMOSOME BANDING AND ASSOCIATED NOMENCLATURE

CHROMOSOME PREPARATION AND CHROMOSOME BANDING METHODS

To study chromosomes under the light microscope, the chromosomes must be suitably condensed—metaphase (or prometaphase) chromosome preparations are required. A peripheral blood sample is taken and separated white blood cells are stimulated to divide by using a mitogen such as phytohemagglutinin. The white blood cells are grown in a rich culture medium containing a spindle-disrupting agent (such as Colcemid) to maximize the number of metaphase cells (cells enter metaphase but cannot progress through the rest of M phase). Prometaphase preparations can also be obtained; they have slightly less-condensed chromosomes, making analysis easier.

Chromosome banding involves treating chromo-some preparations with denaturing agents; alternatively they are digested with enzymes and then exposed to a dye that can bind to DNA. Some dyes preferentially bind to AT-rich sequences; others bind to GC-rich sequences. The dyes show differential binding to different regions across a chromosome that will reflect the relative frequencies of AT and GC base pairs.

The most commonly used method in human chromosome banding is **G-banding**. The chromosomes are treated with trypsin and stained with Giemsa, which preferentially binds AT-rich regions, producing alternating dark bands (Giemsa-positive; AT-rich) and light bands (Giemsa-negative; GC-rich). Because genes are preferentially associated with GC-rich regions, dark bands in G-banding are gene-poor; light bands are gene-rich.

HUMAN CHROMOSOME AND CHROMOSOME BANDING NOMENCLATURE

Human chromosome nomenclature is decided periodically by the International Standing Committee on Human Cytogenetic Nomenclature; see under Further Reading for the most recent ISCN report published in 2020. The nomenclature assigns numbers 1–22 to the autosomes according to perceived size, and uses the symbols p and q to denote, respectively, the short and long arms of a chromosome. Depending on the position of the centromere, chromosomes are described as *metacentric* (centromere at or close to the middle of the chromosome), *submetacentric* (centromere some distance from the middle and from telomeres), or *acrocentric* (centromere close to a telomere).

Each chromosome arm is subdivided into a number of regions, according to consistent and distinct morphological features (depending on the size of the chromosome arm, there may be from one to three regions). Each region is in turn divided into bands, and then into sub-bands and sub-sub-bands, according to the banding resolution (**Figure 1**).

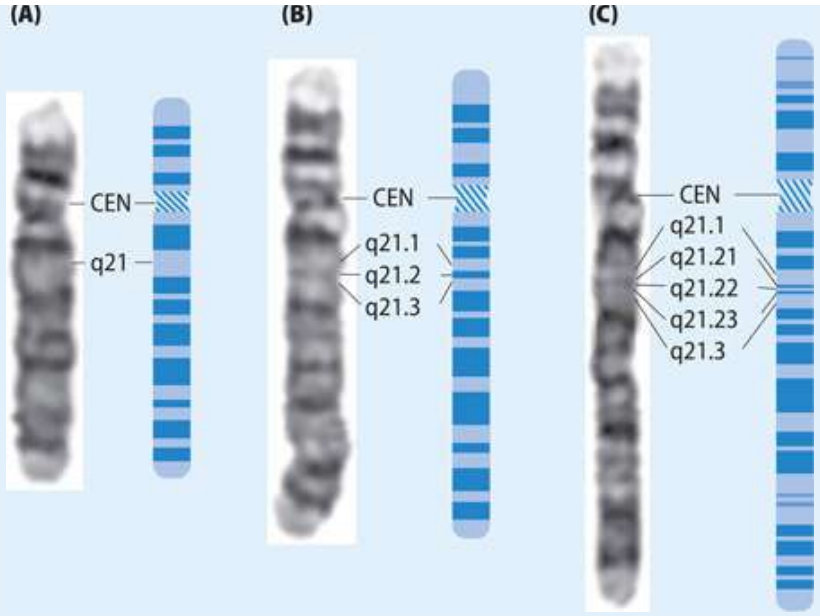


Figure 1 Chromosome banding resolutions can resolve bands, sub-bands and sub-sub-bands. G-banding patterns for human chromosome 4 (with accompanying ideogram to the right) are shown at increasing levels of resolution. The levels correspond approximately to (A) 400, (B) 550, and (C) 850 bands per haploid set, allowing the visual subdivision of bands into sub-bands and sub-sub-bands as the resolution increases. CEN, centromere. For an example of a full set of banded chromosomes, see Figure 2.8. (Adapted from Cross I & Wolstenholme J [2001] in *Human cytogenetics: constitutional analysis*, 3rd ed (Rooney DE, Ed). With permission from Oxford University Press.)

The numbering of regions, bands, sub-bands, and sub-sub-bands is done according to relative proximity to the centromere. If a chromosome arm has three regions, the region closest to the centro-mere would be region 1, and the one closest to the telomere would be region 3. For example, the band illustrated in Figure 1A would be known as 4q21 for these reasons: it is located on the long arm of chromosome 4 (= 4q); it resides on the second (of three regions) on this chromosome arm (= 4q2); within this region it is the nearest band (band 1) to the centro-mere. The last two digits of band 4q21 are therefore pronounced two-one (not twenty-one) to mean region two, band one. Similarly, in Figure 1C the numbers following 4q in the sub-sub-band 4q21.22 are pronounced two-one-point-two-two.

Note also that in chromosome nomenclature the words **proximal** and **distal** are used to indicate the relative position on a chromosome with respect to the centromere. Thus, proximal Xq means the segment of the long arm of the X that is

closest to the centromere. Similarly, distal 3p means the portion of the short arm of chromosome 3 that is most distant from the centro-mere (= closest to the telomere).

According to their distribution in cells of the body, chromosome abnormalities can be classified into two types. A **constitutional** abnormality is present in all nucleated cells of the body and so must have been present very early in development. It can arise as a result of an abnormal sperm or egg, through abnormal fertilization, or through an abnormal event in the very early embryo. A somatic (or acquired) abnormality is present in only certain cells or tissues of a person, who is therefore a genetic **mosaic** (by possessing two populations of cells with altered chromosome or DNA content, each deriving from the same zygote).

Chromosomal abnormalities, whether constitutional or somatic, can be subdivided into two categories: structural abnormalities (which arise through chromosome breakage events that are not repaired), and numerical abnormalities (changes in chromosome number that often arise through errors in chromosome segregation). See **Table 7.10** for a guide to the nomenclature of human chromosome abnormalities.

TABLE 7.10 NOMENCLATURE OF CHROMOSOME ABNORMALITIES

Type of abnormality	Examples	Explanation/notes
NUMERICAL		
Triploidy	69,XXX, 69,XXY, 69,XYY	a type of polyploidy
Trisomy	47,XX,+21	gain of a chromosome is indicated by+
Monosomy	45,X	a type of aneuploidy; loss of an autosome is indicated by +
Mosaicism	47,XXX/46,XX	a type of mixoploidy
STRUCTURAL		
Deletion	46,XY,del(4)(p16.3)	terminal deletion (breakpoint at 4p16.3)
This is a short nomenclature; a more complicated nomenclature is defined by the current ISCN report that allows complete description of any chromosome abnormality; see McGowan-Jordan J et al. (2020) under Further Reading.		

Type of abnormality	Examples	Explanation/notes
	46,XX,del(5)(q13q33)	interstitial deletion (5q13-q33)
Inversion	46,XY,inv(11)(p11p15)	paracentric inversion (breakpoints on same arm)
Duplication	46,XX,dup(1)(q22q25)	duplication of region spanning 1 q22 to 1 q25
Insertion	46,XX,ins(2)(p13q21q31)	a rearrangement of one copy of chromosome 2 by insertion of segment 2q21—q31 into a breakpoint at2p13
Ring chromosome	46,XY,r(7)(p22q36)	joining of broken ends at 7p22 and 7q36
Marker	47,XX,+mar	indicates a cell that contains a ***marker chromosome*** (an extra unidentified chromosome)
Reciprocal translocation	46,XX,t(2;6)(q35;p21.3)	a balanced reciprocal translocation with breakpoints at 2q35 and 6p21.3
Robertsonian translocation (gives rise to one derivative chromosome)	45,XY,der(14;21)(q10;q10)	a balanced carrier of a 14;21 Robertsonian translocation. q10 is not really a chromosome band, but indicates the centromere; der— derivative—is used when one chromosome from a translocation is present
	46,XX,der(14;21)(q10;q10),+21	an individual with Down syndrome possessing one normal chromosome 14, a Robertsonian translocation 14;21 chromosome, and two normal chromosome 21s

This is a short nomenclature; a more complicated nomenclature is defined by the current ISCN report that allows complete description of any chromosome abnormality; see McGowan-Jordan J et al. (2020) under Further Reading.

Structural chromosomal abnormalities

As detailed in Section 4.1, abnormal chromosome breaks (caused by double-strand DNA breaks) occur as a result of unrepaired damage to DNA or through faults in the recombination process. Chromosome breaks that occur during the G2 phase (after the DNA has replicated) are really *chromatid* breaks: they affect only one of the two sister chromatids. The breaks occurring during the G1 phase that are not repaired by S phase (when the DNA replicates) become chromosome breaks (both sister chromatids are affected). A cell with highly damaging chromosome breaks may often be removed by triggering cell death mechanisms; if it survives with unrepaired breaks, chromosomes with structural abnormalities can result.

Errors in recombination that produce structural chromosome abnormalities can occur at meiosis. Paired homologs are normally subjected to recombination mechanisms that ensure the breakage and rejoining of non-sister chromatids, but if recombination occurs between mispaired homologs, the resulting products may have structural abnormalities. Intrachromatid recombination can also be a source of structural abnormalities.

A form of somatic recombination also occurs naturally in B and T cells in which the cellular DNA undergoes programmed rearrangements to make antibodies and T cell receptors. Abnormalities in these recombination processes can also cause structural chromosomal abnormalities that may be associated with cancer (described in Chapter 10).

Structural chromosome abnormalities are often the result of incorrect joining together of two broken chromosome ends. Different mechanisms are possible, as detailed in the following subsections.

Microdeletions and microduplications

As detailed in Section 7.3, exchanges between mispaired repeats on opposing chromatids, or even in the same chromatid can produce duplications and deletions several Mb long within a chromosome arm. Such microdeletions and microduplications result in simultaneous change in copy number of genes in the affected region, and clinical phenotypes can result if any of these genes are especially *dosage-sensitive* (just as they do in the case of whole chromosome duplication, as in trisomy 21).

Microduplications can cause disease by increasing the copy number of a single dosage-sensitive gene. For example, Charcot-Marie-Tooth syndrome 1A (OMIM 118220) can be caused by duplications at 17p12 that include the dosage-sensitive

PMP22 gene (which makes peripheral myelin protein 22). Having three copies of *PMP22* instead of the normal two copies is sufficient to cause problems for cells, as are activating point mutations in *PMP22* that cause overexpression. Microdeletions can cause disease in different ways, as shown by the examples listed in **Table 7.11**.

TABLE 7.11 SOME OF THE DIFFERENT WAYS IN WHICH CHROMOSOMAL MICRODELETIONS CAUSE DISEASE

Cause of disease	Examples	Key disease gene(s) / allele
Reduced copy number of a single dosage-sensitive gene	Deletions at 16p13.3 associated with Rubinstein-Taybi syndrome (OMIM 180849)	dosage-sensitive *CBP* gene
	Deletions at 20p12.1 associated with Alagille syndrome type 1 (OMIM 118450)	dosage-sensitive *JAG1* gene
Reduced copy number of >1 dosage-sensitive gene *(segmental aneuploidy)*	Deletions at 11 p13 associated with WAGR (Wilms tumor, aniridia, genitourinary abnormalities and mental retardation—OMIM 194072)	dosage-sensitive *PAX9, WT1* genes
Loss of the active allele of one or more imprinted genes	Deletions of maternal 15q 11 -q 13 associated with Angelman syndrome (OMIM 105830)	maternal *UBE3A* allele (the only active allele)
	Deletions of paternal 15q 11 -q 13 associated with Prader-Willi syndrome (OMIM 176270)	paternal *SNRPN* allele (the only active allele)
Loss of the only allele for genes on the male X-chromosome	Contiguous gene syndrome causing Duchenne muscular dystrophy, chronic granulomatous disease and retinitis pigmentosa (PMID 4039107)	DMD, CYBB, RPGR (single alleles in males due to hemizygosity)

Large-scale duplications, deletions, and inversions

Still larger changes can occur when breaks occur in both arms of a chromo-some. If a single chromosome sustains two breaks, incorrect joining of fragments can result in chromosome material being lost (deletion), switched round in the reverse direction (inversion), or included in a circular chromosome (a ring chromosome) (Figure 7.12).

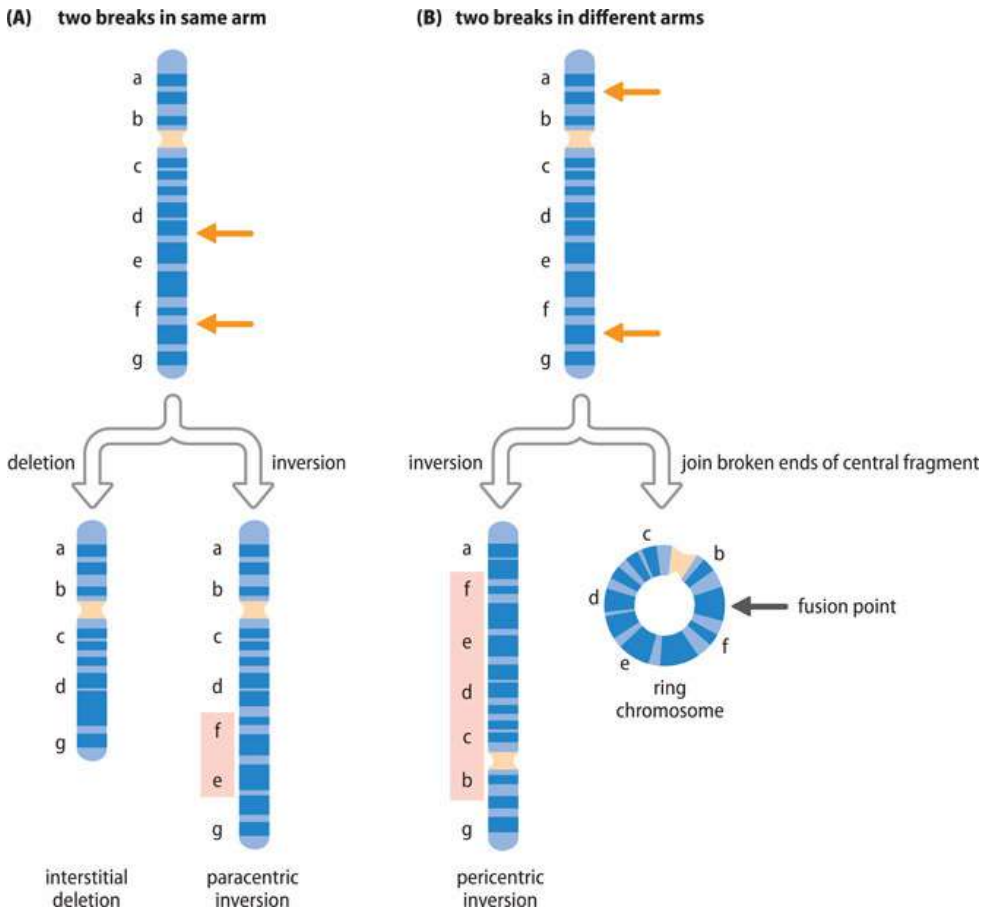


Figure 7.12 Stable outcomes after incorrect repair of two breaks on a single chromosome. (A) Incorrect repair of two breaks (orange arrows) occurring in the same chromosome arm can involve loss of the central fragment (here containing hypothetical regions e and f) and rejoining of the terminal fragments (deletion), or inversion of the central fragment through 180° and rejoining of the ends to the terminal fragments (called a paracentric inversion because it does not involve the centromere). (B) When two breaks occur on different arms of the same chromosome, the central fragment (encompassing hypothetical regions b to f in this example) may invert and rejoin the terminal fragments (pericentric inversion). Alternatively, because the central fragment contains a centromere, the two ends can be joined to form a stable ring chromosome, while the acentric distal fragments are lost. Like other repaired chromosomes that retain a centromere, ring chromosomes can be stably propagated to daughter cells.

Structurally abnormal chromosomes with a single centromere can be stably propagated through successive rounds of mitosis. However, any repaired chromosome that lacks a centromere (an *acentric* chromosome) or possesses two centromeres (a *dicentric* chromosome) will normally not segregate stably at mitosis, and will eventually be lost.

Chromosomal translocations

If two different chromosomes each sustain a single break, incorrect joining of the broken ends can result in the movement of chromosome material between chromosomes (**translocation**). A *reciprocal translocation* is the general term used to describe an exchange of fragments between two chromosomes (Figure 7.13A). If an acentric fragment from one chromosome (one that lacks a centromere) is exchanged for an acentric fragment from another, the products each have a centromere and are stable in mitosis. Structurally rearranged chromosomes like this that have a centromere are known as **derivative chromosomes**. Exchange of an acentric fragment for a centric fragment results in acentric and dicentric chromosomes that are normally unstable in mitosis (but see below for an exceptional class of translocations in which dicentric products are stable).

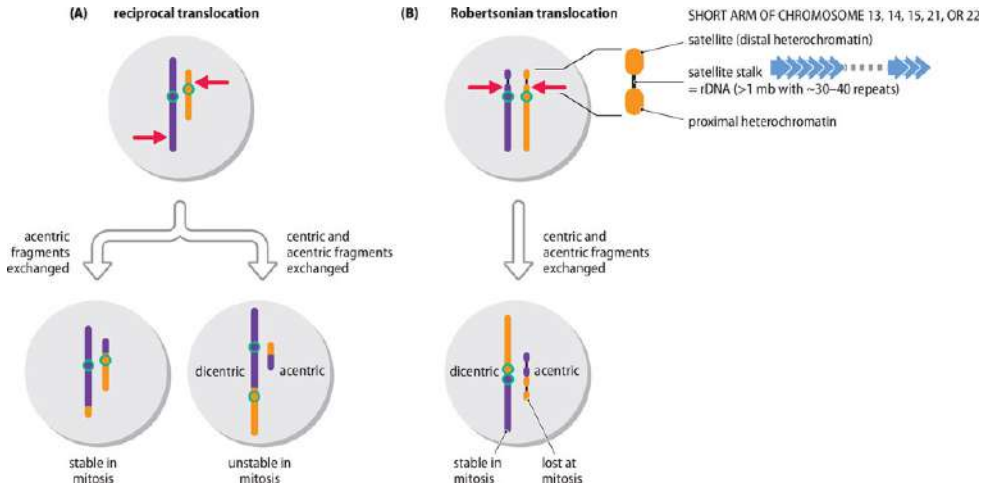


Figure 7.13 Reciprocal and Robertsonian translocations. (A) Reciprocal translocation. The *derivative* chromosomes produced by the translocation are stable in mitosis when one acentric fragment is exchanged for another, but when a centric fragment is exchanged for an acentric fragment, the derivative chromosomes are usually unstable. (B) Robertsonian translocation (centric fusion). This is a highly specialized reciprocal translocation in which exchange of centric and acentric fragments produces a dicentric chromosome that is nevertheless stable in mitosis. It occurs

exclusively after breaks in the short arms of two of the acrocentric chromosomes 13, 14, 15, 21, and 22. As illustrated, the short arms of the human acrocentric chromosomes have a common structure. A region of distal heterochromatin (called a *satellite)* is joined to a proximal heterochromatic region by a thin *satellite stalk* made up of *ribosomal DNA* (rDNA), an array of tandem DNA repeats that each make three types of rRNA. Breaks that occur close to the centromere can result in a dicentric chromosome in which the two centromeres are so close that they can function as a single centromere. The loss of the small acentric fragment has no phenotypic consequences.

For a chromosomal translocation to occur, the regions containing double-strand DNA breaks in the two participating chromosomes must be in close proximity (allowing incorrect joining prior to double-strand break repair). The spatial distribution of chromosomes in the nucleus is not random, and chromosomes tend to occupy certain "territories". Chromosomes that tend to be physically closer to each other are more likely to engage in translocation with each other. For example, human chromosomes 4, 13, and 18 are preferentially located at the periphery of the nucleus and frequently translocate with each other but not with physically distant chromosomes localized in the interior of the nucleus. Specific types of translocations are common in certain cancers, and may reflect close physical association of the two chromosomal regions that participate in translocation.

One exceptional form of chromosome association occurs between the very small short arms of the five human acrocentric chromosomes (chromosomes 13, 14, 15, 21, and 22). Each of the short arms of these chromosomes has about 30–40 large tandem DNA repeats, each containing sequences for making three ribosomal RNAs: 28S, 18S, and 5.8S rRNAs; the five ribosomal DNA (rDNA) regions congregate at the nucleolus to produce these rRNAs. The close physical association of these five chromosome arms is responsible for a specialized type of translocation, called Robertsonian translocation or centric fusion, that involves breaks in the short arms of two different acrocentric chromosomes followed by exchange of acentric and centric fragments to give acentric and dicentric products (Figure 7.13B).

The acentric chromosome produced by a Robertsonian translocation is lost at mitosis without consequence (it contains just highly repetitive noncoding DNA plus rRNA genes that are also present at high copy number on the other acrocentric chromosomes). The other product is an unusual dicentric chromosome that is stable in mitosis: the two centromeres are in close proximity (centric fusion) and often function as one large centromere so that the chromosome segregates regularly. (Nevertheless, such a chromosome may present problems during gametogenesis.)

More complex translocations can involve multiple chromosome breakages. Insertions typically require at least three breaks: fragment liberated by two breaks in one chromosome arm inserts into another break located in another region of the same chromosome or in a different chromosome.

Isochromosomes

An additional rare class of structural abnormality is a symmetrical **isochromo-some** in which the arms of the chromosome are mirror images of one another, having either two long arms or two short arms (Figure 7.14). The centromere *appears* to divide transversely instead of longitudinally. Isochromosomes can form by a type of misdivision of the centromere but more commonly arise through the breakage and fusion of sister chromatids. Overall the effect equates to a combined deletion-duplication event (deletion of one chromosome arm and duplication of the other). Human isochromosomes are rare, except for i(Xq) and also i(21q), an occasional contributor to Down syndrome.

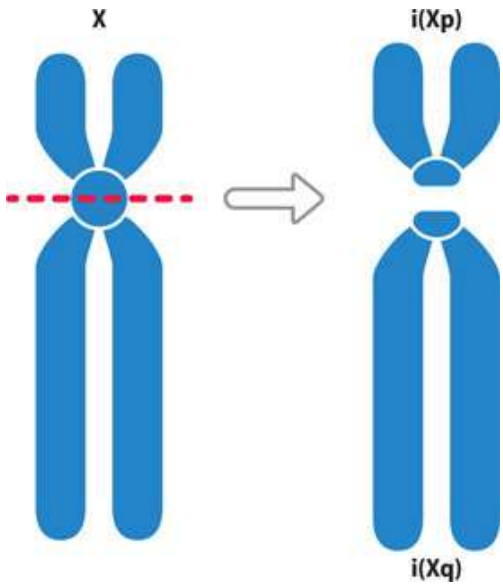


Figure 7.14. Isochromosome formation. This can sometimes occur by a type of misdivision of the centromere as shown here (for alternative origins, see text).

Chromosomal abnormalities involving gain or loss of complete chromosomes

Three classes of numerical chromosomal abnormalities can be distinguished: polyploidy and aneuploidy (summarized in Table 7.12), plus mixoploidy (described below).

TABLE 7.12 CLINICAL CONSEQUENCES OF NUMERICAL CHROMOSOME ABNORMALITIES

Abnormality	Clinical consequences
POLYPLOIDY	
Triploidy (69,XXX or 69,XYY)	1 -3 % of all conceptions; almost never born live and do not survive long
ANEUPLOIDY (AUTOSOMES)	
Nullisomy (missing a pair of homologs)	lethal at pre-implantation stage of embryonic development
Monosomy (one chromosome missing)	lethal during embryonic development
Trisomy (one extra chromosome)	usually lethal during embryonic* orfetal* stages, but individuals with trisomy 13 (Patau syndrome) and trisomy 18 (Edwards syndrome) may survive to term; those with trisomy 21 (Down syndrome) may survive beyond age 40
ANEUPLOIDY (SEX CHROMOSOMES)	
Additional sex chromosomes	individuals with 47,XXX, 47,XXY, and 47,XYY all experience relatively minor problems and a normal lifespan
Lacking a sex chromosome	while 45,Y is never viable, in 45,X (Turner syndrome), about 99 % of cases abort spontaneously; survivors are of normal intelligence but are infertile and show minor physical diagnostic characteristics

* In humans, the embryonic period spans fertilization to the end of the eighth week of development; fetal development then begins and lasts until birth.

Polyploidy

Three per cent of recognized human pregnancies produce a triploid embryo (**Figure 7.15A**). The usual cause is two sperm fertilizing a single egg (dispermy), but triploidy is sometimes attributable to fertilization involving a diploid gamete. With three copies of every autosome, the dosage of autosomal genes might be expected to be balanced, but triploids very seldom survive to term, and the condition is not compatible with life (but diploid/triploid mosaics can survive). The lethality in triploids may be due to an imbalance between products encoded on the X chromosome and autosomes, for which X-chromosome inactivation would be unable to compensate.

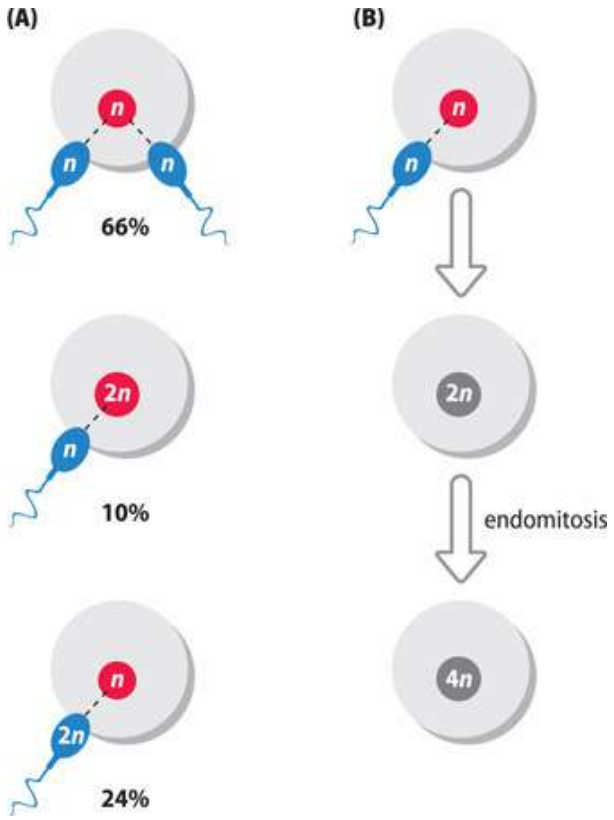


Figure 7.15 Origins of triploidy and tetraploidy. (A) Origins of human triploidy. Dispermy (top) is the principal cause, accounting for 66 % of cases. Triploidy is also caused by diploid gametes that arise by occasional faults in meiosis, such as nondisjunction (see Figure 7.16); fertilization of a diploid ovum (middle) and fertilization by a diploid sperm (bottom) account for 10 % and 24 % of cases, respectively. (B) Tetraploidy involves normal fertilization and fusion of gametes to give a normal zygote. Subsequently, however, tetraploidy arises when DNA replicates without subsequent cell division (*endomitosis*).

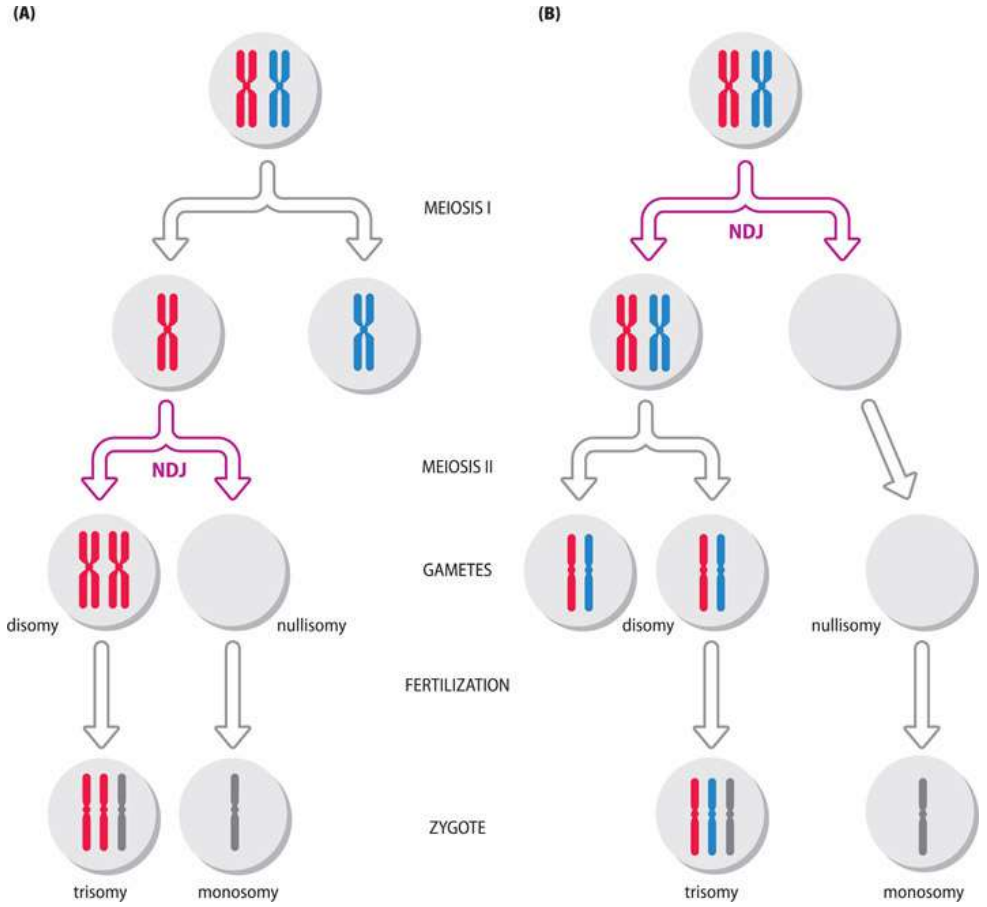


Figure 7.16 Meiotic nondisjunction and its consequences. In meiotic nondisjunction (NDJ) a pair of homologous chromosomes fail to disjoin and migrate to the same daughter cell, instead of being individually allocated to each daughter cell. That can happen at the second meiotic division (A) or at the first meiotic division (B). The result is that a person produces abnormal gametes, with either no copies of a chromosome (nullisomy) or two copies of a chromosome (disomy). In the latter case, the two chromosome copies may originate from a single parent (for NDJ at meiosis I) or individually from each parent (for NDJ at meiosis I). Fertilization with a normal gamete containing one copy of the chromosome (shown in gray) results in trisomy or monosomy.

Tetraploidy (Figure 7.15B) is much rarer and always lethal. It is usually due to failure to complete the first zygotic division: the DNA has replicated to give a content of 4C, but cell division has not then taken place as normal. Although constitutional polyploidy is rare and lethal, some types of cell are naturally polyploid in all normal individuals—for example, our muscle fibers are formed by recurrent cell fusions that result in multinucleate syncytial cells.

Aneuploidy

Normally our nucleated cells have complete chromosome sets (euploidy), but sometimes one or more individual chromosomes are present in an extra copy or are missing (**aneuploidy**). In trisomy, three copies of a particular chromo-some are present in an otherwise diploid cell, such as trisomy 21 (47,XX,+21 or 47,XY,+21) in Down syndrome. In monosomy a chromosome is lacking from an otherwise diploid state, as in monosomy X (45,X) in Turner syndrome. Cancer cells often show extreme aneuploidy, with multiple chromosomal abnormalities.

Aneuploid cells arise through two main mechanisms. One is **nondisjunction**, in which paired chromosomes fail to separate (disjoin) during meiotic anaphase I and migrate to the same daughter cell, or sister chromatids fail to disjoin at either meiosis II or mitosis. Nondisjunction during meiosis produces gametes with either 22 or 24 chromosomes, which after fertilization with a normal gamete produce a trisomic or monosomic zygote (**Figure 7.16**). If nondis-junction occurs during mitosis, the individual is a mosaic with a mix of normal and aneuploid cells.

Aneuploidy can also occur by **anaphase lag**. If, during anaphase, a chromo-some or chromatid is delayed in its movement and lags behind the others, it may not be incorporated into one of the daughter nuclei. Chromosomes that do not enter a daughter cell nucleus are eventually degraded.

Having the wrong number of chromosomes has serious, usually lethal, consequences (Table 7.12). Even though the extra chromosome 21 in a person with trisomy 21 (Down syndrome) is a perfectly normal chromosome, inherited from a normal parent, its presence causes multiple abnormalities that are present from birth (congenital). Embryos with trisomy 13 or trisomy 18 can also survive to term but both result in severe developmental malformations, respectively Patau syndrome and Edwards syndrome. Other autosomal trisomies are not compatible with life. Autosomal monosomies have even more catastrophic consequences than trisomies and are invariably lethal at the earliest stages of embryonic life. We consider in Section 7.7 how gene dosage problems in aneuploidies result in disease.

Maternal age effects in Down syndrome

In principle, the nondisjunction that causes a gamete to have an extra copy of chromosome 21 could occur at either meiotic division in spermatogenesis or oogenesis, but in practice in about 70 % of cases it occurs at meiosis I in the mother.

This is almost certainly a consequence of the extremely long duration of meiosis I in females (it begins in the third month of fetal life but is arrested and not completed until after ovulation). That means this one meiotic division can take decades; by contrast, male meiosis occurs continuously in the testes from puberty to old age. Not only is there a sex difference in the origin of the extra chromosome 21 but there is also a very significant *maternal age effect*. Thus the risk of having a Down syndrome child in a 20-year-old pregnant woman is about 1 in 1500, but that increases to about 1 in 25 for a 45-year-old woman.

Mixoploidy: mosaicism and chimerism

Mixoploidy means having two or more genetically different cell lineages within one individual. Usually, the different cell populations arise from the same zygote (*mosaicism*). More rarely, a person can have different cell populations that originate from different zygotes and is described as a **chimera**; spontaneous chimerism usually arises by the aggregation of fraternal twin zygotes or immediate descendant cells within the very early embryo. Abnormalities that would otherwise be lethal (such as triploidy) may not be lethal in mixoploid individuals.

Aneuploidy mosaics, with a proportion of normal cells and a proportion of aneuploid cells, are common. This type of mosaicism can result when non-disjunction or chromosome lag occurs in one of the mitotic divisions of the early embryo (any monosomic cells that are formed usually die). Polyploidy mosaics (such as human diploid/triploid mosaics) are occasionally found. As the gain or loss of a haploid set of chromosomes by mitotic nondisjunction is extremely unlikely, human diploid/triploid mosaics most probably arise by fusion of the second polar body with one of the cleavage nuclei of a normal diploid zygote.

7.6 MOLECULAR PATHOLOGY OF MITOCHONDRIAL DISORDERS

Up until now we have focused on disease resulting from changes to the nuclear DNA where the vast majority (99.9 %) of our genes reside. The very small (16.6 kb) circular mitochondrial genome is thought to have originated after an anerobic proto-eukaryotic cell engulfed an aerobic proteobacterium in a type of co-operative symbiosis about two billion years ago, when the amount of oxygen in the

atmosphere was increasing rapidly. The internalized cell (endosymbiont) was not destroyed; instead, its genome and protein synthesis capacity were retained.

Over time, the genome of the engulfing cell expanded to become the large nuclear genome. Many of the sequences of the internalized cell's genome—and their functions—were inserted into the nucleus, and the genome of the endosymbiont became depleted (through genetic redundancy and sequence loss). Nuclear insertion of mtDNA sequences from degraded mitochondria is an evolutionarily ongoing process that continues to this day. It has resulted in a family of defective mtDNA sequences scattered across the nuclear genome, previously called mitochondrial pseudogenes but now known as NUMT (**nu**clear-**mit**ochondrial) sequences. Some NUMTs have very high sequence similarity to regions of mtDNA and may hinder mtDNA analyses.

Our mitochondrial genome has a total of 37 RNA genes, producing 22 mitochondrial tRNAs and two mitochondrial rRNAs. It makes just 13 proteins, all components of the multisubunit protein complexes needed for oxidative phosphorylation. However, the close to 140 other proteins needed for oxidative phosphorylation (plus another >1100 proteins needed for mitochondrial function) are made by nuclear genes, translated on cytoplasmic ribosomes and imported into mitochondria. Because mitochondria act as a cell's power source, "mitochondrial disease" is the generic term for a clinically heterogeneous group of disorders that directly affect oxidative phosphorylation. Multiple systems may be involved, but the organs and tissues most often affected are usually those with high energy demands.

Mitochondrial disorders arising from mutation of nuclear genes show Mendelian inheritance: autosomal dominant, autosomal recessive and X-linked forms are all seen (plus occasional *de novo* mutations). Although there are just 37 genes in mtDNA, they all play vital roles in energy production in our cells. As a result, defects in mitochondrial genes cause, or contribute to, a very wide variety of genetic conditions.

In this section we consider disorders resulting from mtDNA variants of strong effect. This clinically diverse grouping includes many disorders that result from large deletions (often affecting multiple genes), or to point mutations. The disorders can be genetically heterogeneous: a single disorder may arise from mutations in any one of several possible genes in mtDNA, and in some cases mutations in nuclear genes may also produce the same phenotype. Because of some peculiar features of mitochondria, disorders due to mutations in mtDNA show some unique properties. Pathogenic mtDNA variants can also make important contributions to complex genetic diseases, as described in Chapter 8.

Mitochondrial disorders due to mtDNA mutation show maternal inheritance and variable proportions of mutant genotypes

During fertilization, a sperm injects its nuclear DNA into an oocyte. Paternal mtDNA is not normally transferred; even if it were, it would be recognized as being foreign and be degraded within the fertilized oocyte. As a result, mitochondrial genes appear to be exclusively maternally inherited—but NUMTs are inherited from both parents.

An important difference between nuclear DNA and mtDNA is the control of DNA replication and segregation of genetic material into daughter cells. For nuclear DNA sequences, DNA replication is very tightly constrained to ensure precise doubling of the amount of genetic material. And segregation of the genetic material into daughter cells is tightly controlled to ensure equal division of maternal and paternal sequences into the two daughter cells. As a result, if a person inherits a normal allele of gene *X* from one parent and a mutant allele from the other parent, the ratio of normal allele to mutant allele will be uniformly 1:1 in all diploid cells (barring any copy number changes), and gametes would contain either a normal or a mutant allele.

Unlike the nuclear genome, mtDNA turns over continuously and independently of the cell cycle, in both dividing and non-dividing cells (such as muscle cells and neurons). The mtDNA undergoes a type of *relaxed replication*: individual mtDNA molecules replicate at random, making sometimes multiple copies at a time (*clonal expansion*). Overall, an approximately constant total number of mtDNA molecules is maintained in the cell (but the total number per cells can vary according to the cell type, somewhere in a range of from 1000–10 000 copies in most cell types, and >100 000 copies in an egg cell). Different mtDNA variants can co-exist in the cells of a person—a situation known as **heteroplasmy**. But unlike for nuclear DNA, the ratio of a normal mtDNA sequence and a mutant one is not fixed, and can be unpredictably variable for different reasons, as listed below.

- *Differential replication of mutant and normal mtDNA*. From studies of model organisms we know that some mutant mtDNA sequences with large mtDNA deletions or point mutations in the major control sequence are able to replicate more rapidly than normal mtDNA copies. The proportion of mutant mtDNA in cells can then increase over time.
- *Vegetative segregation*. Unlike nuclear DNA, mtDNA molecules in a dividing cell are often segregated unequally to the daughter cells. The

dividing cell splits unequally so that one daughter cell acquires the majority of both the cytoplasm and the mitochondria.

- *Genetic bottleneck in female germ line development.* Before giving rise to an egg cell, primary human oocytes originate from diploid primordial germ cells (PGCs) through a series of mitotic divisions. In very early PGCs a genetic bottleneck occurs whereby only a very small number of mitochondria (and mtDNA molecules) are transmitted, randomly, to daughter cells; as a result, the resulting cells can have very different proportions of mutant mtDNA. Subsequently, a large increase in mtDNA copy number occurs (to give ultimately >100 000 mtDNA copies in egg cells), but eggs from a woman can continue to show wide differences in the proportion of mutant mtDNA (Figure 7.17).

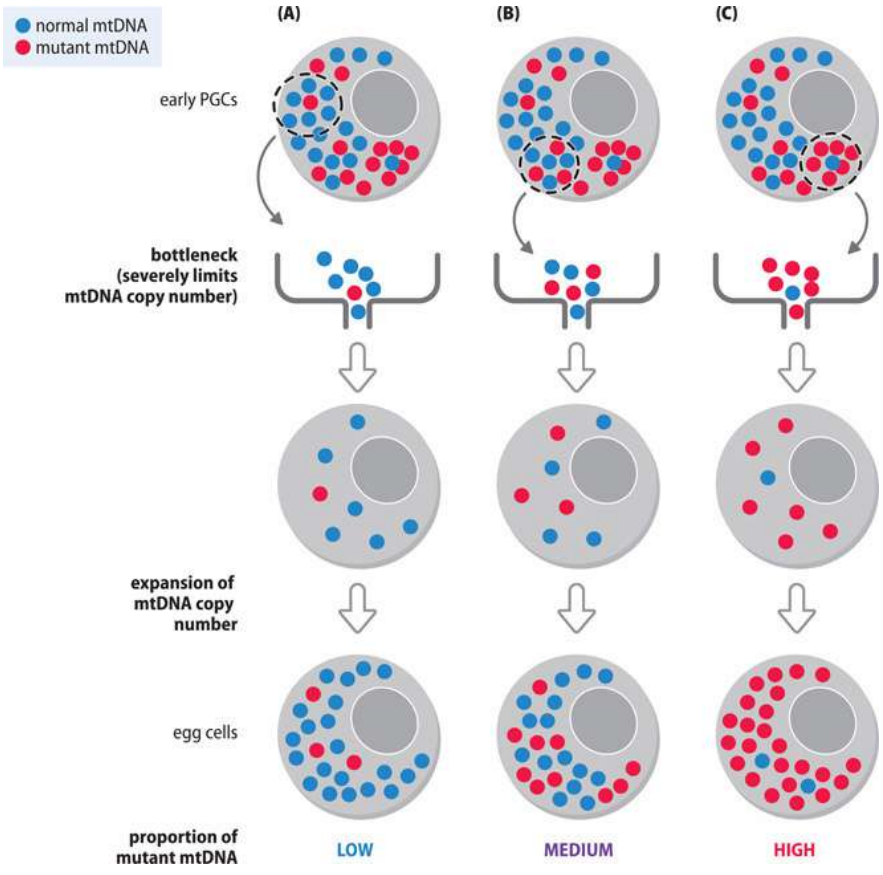


Figure 7.17. The mitochondrial genetic bottleneck: how a heteroplasmic woman can give rise to egg cells with quite different mutation loads. Mammalian egg cells are derived by two successive meiotic divisions from primary oocytes (not shown) that in turn originate by sequential mitotic divisions from primordial germ cells (PGCs). In this example we consider how eggs might be produced from a mother who is approximately

50 % heteroplasmic for a pathogenic mtDNA mutation. The early PGCs might show 50 % heteroplasmy, but then because of the mitochondrial genetic bottleneck a very limited number of the mtDNAs in these cells are passed to daughter cells. Depending on which mtDNAs are transmitted through the bottleneck, later PGCs may have different proportions of mutant mtDNA, low (A), intermediate (B) or high (C). After the subsequent large increase in mtDNA copy numbers, such mutant frequencies may persist in the more mature egg cells.

The variable heteroplasmy in maternal eggs illustrated in Figure 7.17 poses problems for genetic counselling. A woman affected by a mtDNA disorder has generally high levels of mutant mtDNA and because the disorder is maternally transmitted, there is a risk of an affected child every time she conceives. However, the different egg cells that she produces can show a wide range of heteroplasmy from small proportions of mutant mtDNA to high mutation loads. As we explore in the next section, the development of a disease phenotype is dependent on a threshold value of heteroplasmy that can vary according to the type of mitochondrial disorder.

The two major classes of pathogenic DNA variant in mtDNA: large deletions and point mutations

Associated pathogenic mutations have been identified in virtually all the genes in the mitochondrial genome. Although the frequency of pathogenic mtDNA mutations in the population is high (about 1 in 200 people), the majority of people carrying them may be unaffected (because of low proportions of mutant mtDNA; note that common mtDNA variants of weak effect may also play a role in common genetic disorders, such as Parkinson disease). Here, we focus on rare mtDNA variants of strong effect that come in two common types: those affecting single genes through point mutations, and large deletions that typically affect multiple genes. Because of variable heteroplasmy, the mutation load in cells and tissues can vary. When mutant mtDNAs reach a certain proportion of the total mtDNA, a *threshold* level, a biochemical phenotype resulting from inefficient oxidative phosphorylation results and clinical symptoms develop. As we show below, threshold heteroplasmy values can vary according to the nature of the mutation.

The connection between large mtDNA deletions and short repeats

Large-scale mtDNA deletions are frequent, and the region spanning nucleotide positions 5 700 to 16 500, almost two-thirds of the mitochondrial genome, has been viewed as a high frequency deletion zone. The deletions can be large—often from 4 kb to approaching 8 kb in length.

The high frequency of large deletions may be yet another case of repeat sequences predisposing to pathogenesis (previously summarized in Table 7.1). Examination of large pathogenic mtDNA deletions reveals that in many cases the sequence deleted in mtDNA is normally flanked by short, often perfect, direct repeats (see **Table 7.13** for examples). Physical association of the direct repeats may predispose to deletions, either during mtDNA replication (when the single-stranded repeats are exposed), or more likely during repair of double-strand breaks. Some mtDNAs with large deletions can replicate more rapidly than normal mtDNA copies. Large deletions may result in certain clinical phenotypes, as listed below:

- Kearns-Sayre syndrome: a multisystem disorder with onset before 20 years of age with progressive external ophthalmoplegia (PEO) and often pigmentary retinopathy; additional features may include cerebellar ataxia, impaired growth, hypoadrenalism, diabetes mellitus and cardiomyopathy and occasionally sensorineural hearing loss and cognitive impairment.
- Around one-third of cases have the common 4.977 kb deletion (see Table 7.13).
- Pearson syndrome: sideroblastic anemia and pancreas deficiency (commonly exocrine pancreas deficiency, but endocrine deficiency also often involved). Frequently a devastatingly fatal condition with profound anemia, thrombocytopenia, and lactic acidosis in the neonatal-infantile period.
- PEO: ptosis, impaired eye movements, oropharyngeal weakness, and variably severe proximal limb weakness.

Mitochondrial deletion disorders are rarely inherited: a best estimate is a 1 in 24 risk of inheritance. Usually, they originate by *de novo* deletion in the maternal germ line or in the very early embryo. Often, there is a quite low threshold of ~50 % to 60 % heteroplasmy—if the proportion of mutant mtDNA exceeds this level the bioenergetic capacity of tissues declines to a level where symptoms are evident. Interested readers can find a detailed review on mitochondrial deletion disorders at https://www.ncbi.nlm.nih.gov/books/NBK1203/)

TABLE 7.13 DIRECT REPEATS ARE HOTSPOTS FOR PATHOGENIC DELETIONS IN mtDNA *

Deletion size	Sequence and location of repeats	
	Repeat 1	Repeat 2
4.420 kb	10942 - 10951 AACAACCCCC	15362 - 15371 AACAACCCCC
4.977 kb	8470 - 8482 ACCTCCCTCACCA	13447 - 13459 ACCTCCCTCACCA
7.521 kb	7 975 - 7 982 AGGCGACC	15496 - 15503 AGGCGACC
7.664 kb	6325 - 6341 CCTCC**G**TAGACCTAACC	13989 - 14004 CCTCCT**A**GACCTAACC
7.723 kb	6076 - 6084 TCACAGCCC	11964 - 11972 TCACAGCCC

* Numbers adjacent to sequences are nucleotide co-ordinates in mtDNA—see Figure 2.12 on page 45 for the mtDNA gene map. The pairs of direct repeats are identical except for a 1 bp difference in those causing the 7.664kb deletion. The 4.977 kb deletion is particularly common—see text.

Mitochondrial disorders arising from mtDNA point mutations

Pathogenic point mutations have been reported in virtually all genes in mtDNA. Whereas genetic redundancy offers protection against mutation in nuclear rRNA and tRNA genes (which are all present in multiple copies), the single-copy rRNA and tRNA genes in mtDNA are more exposed to harmful mutations, and pathogenic mutations have been identified in the *MT-RNR1* gene (making 12S rRNA), and all mitochondrial tRNA and protein-coding genes. Note, however, that as of January 2021, the MITOMAP database had confirmed pathogenic variants in 17 of the 22 tRNA genes.

Most of the associated clinical phenotypes are heteroplasmic, with varying thresholds beyond which biochemical and clinical phenotypes manifest. Most pathogenic mutations in mitochondrial tRNA genes show thresholds of >90 % heteroplasmy. Some mitochondrial disorders, notably Leber hereditary optic neuropathy, are said to be homoplasmic: the proportion of mutant mtDNA appears to be virtually 100 %.

Various mtDNA disorders show very considerable genetic heterogeneity, such as Leber hereditary optic neuropathy (LHON) and Leigh syndrome. The former disorder typically presents in young adults as bilateral, painless, subacute visual failure and the great majority of those who lose their vision do so before 50 years of age. Leigh syndrome shows a similar phenotype. A case study, presented in **Clinical Box 7**, shows how investigation of a mtDNA disorder initially believed to be LHON,

revealed the cause to be a pathogenic point mutation associated with Leigh syndrome.

CLINICAL BOX 7 A CASE STUDY: LEIGH SYNDROME

Joanne, a healthy woman aged 41, was about to undergo her first round of IVF when her sister Carole told her that her grandson Archie had been diagnosed with Leigh syndrome (OMIM #256000), a genetically heterogeneous disorder (see Figure 1 for the family tree). Joanne was advised by family members that the Leigh syndrome in her family was maternally inherited, and that she should postpone her plans to start a family until she had been tested. Joanne's mother was in her 70s and generally well, but Carole had developed severe visual impairment at four years of age and remained registered blind.

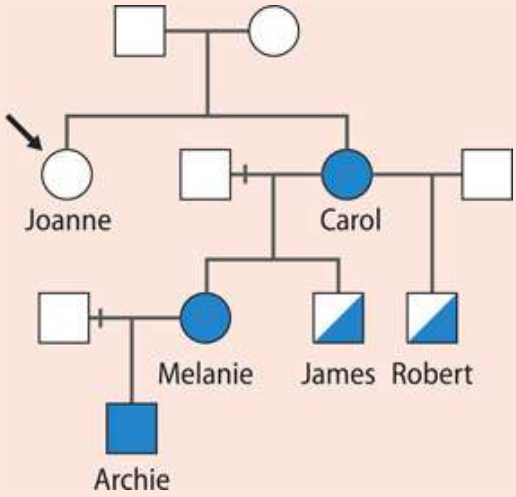


Figure 1 Family tree.

Carole's daughter, Melanie, had also developed profound visual impairment (at age 23), presenting to the ophthalmology department with subacute onset of asymmetric painless central visual loss that became bilateral within days. Fundoscopy had revealed a pattern of optic disc pallor compatible with that observed in Leber hereditary optic neuropathy (LHON; OMIM 535000). However, investigations of the three common mtDNA variants associated with LHON (m.3460G>A, m.11778G>A and m.14484T>C, located respectively in the *MT-ND1, MT-ND4* and *MT-ND6* genes) had been negative.

Follow-up whole mtDNA sequencing revealed a pathogenic variant, m.13051G>A, known to be associated with Leigh syndrome, rather than LHON.

This variant, occurring in the *MT-ND5* gene (which makes the NADH:ubiquinone oxidoreductase core subunit 5 protein), was present at 92 % heteroplasmy (that is, 92 % of the mtDNA molecules had the pathogenic variant; 8 % were normal). Melanie became pregnant at the age of 24 years, just a month before she received this genetic information. Melanie's brothers, James and Robert, remained well and both refused formal testing, but were counselled regarding maternal inheritance of this disease.

Melanie had a normal pregnancy and delivered her son Archie at 40 weeks and 5 days by caesarean section following a failed induction of labor. No resuscitation was required and Archie's early developmental milestones were considered normal, though he didn't walk independently until 17 months and remained unsteady with a broad-based gait when assessed at 2.5 years. Acute onset of strabismus and visual impairment at the age of 2.5 years, in the context of continued gait instability prompted further investigation of Archie including a cranial MRI scan. Bilateral, symmetrical foci of hypoin-tense T1 and hyperintense T2 signal in the nigrostriatal pathways of the upper brainstem and floor of the 4th ventricle (**Figure 2**) raised the clinical suspicion of Leigh syndrome and, given Melanie's diagnosis, rapid analysis for the m.13051G>A variant was requested in blood DNA from Archie. This genetic testing confirmed a heteroplasmy of 96 %. Archie remains visually impaired with optic disc pallor and ataxia, but his developmental trajectory is comparable with that of his peers at present.

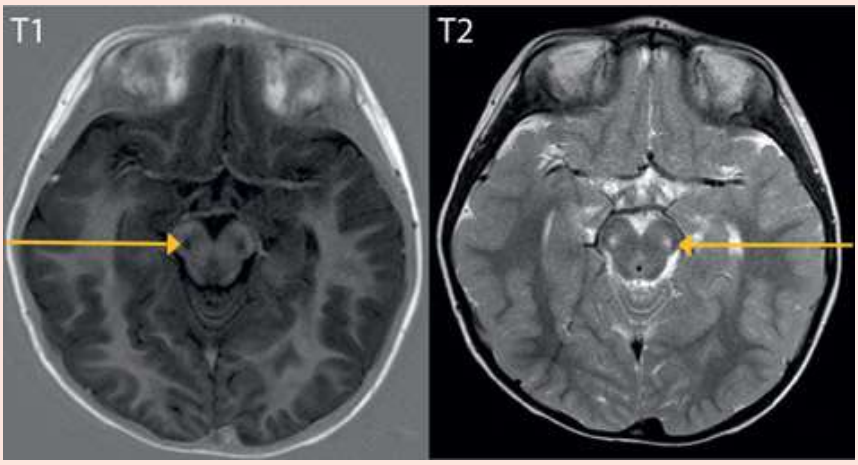


Figure 2 Cranial MRI scan of Archie's optic disc. Yellow arrows indicate the hypointense T1/hyperintense T2 signal change in the substantia nigra of the midbrain.

Further genetic testing within the family confirmed that Carole also harbored the m.13051G>A pathogenic variant but heteroplasmy in blood and urinary sediment

DNA was considerably lower than that recorded in Melanie and Archie. Much to her relief, Joanne tested negative for m.13051G>A in blood and urinary sediment. The possibility that this pathogenic variant had arisen spontaneously in Carole and had been maternally transmitted thereafter could not be investigated because the mother of Joanne and Carole was not available for testing. Joanne has now resumed her plans for IVF.

The optic disc appearance and associated central scotoma that gave rise to the clinical diagnosis of LHON in Melanie, and her mother Carole, has been observed with a number of mtDNA pathogenic variants affecting the activity of complex I of the mitochondrial respiratory chain. Unlike LHON, the clinical features resulting from these variants in *MT-ND5* often extend beyond the optic nerve to involve the brain and other organ systems. Generally, the extent and severity of clinical features in mtDNA-related mitochondrial disease correlate with heteroplasmy—high levels being associated with earlier onset of more severe disease—though the threshold for specific cellular and organ dysfunction may be high and will vary between individuals and between different pathogenic variants. It is also clear that heteroplasmy is not alone in determining pheno-type and other epigenetic and nuclear genetic factors exert substantial influence on outcome. Smoking and excessive alcohol consumption are environmental factors associated with clinical presentation of LHON in those individuals harboring the three common mtDNA variants. Given the preponderance of males affected by LHON (typically a 5:1 ratio for affected males to females) an X-linked nuclear modifier has long been suspected, but not yet confirmed.

7.7 EFFECTS ON THE PHENOTYPE OF PATHOGENIC VARIANTS IN NUCLEAR DNA

Sections 7.2 to 7.5 described how disease-causing genetic changes arise in DNA and chromosomes, and Section 7.6 covered the special case of pathogenesis due to mutations in mtDNA. Here we consider the effects on the phenotype of pathogenic variants in nuclear DNA. There are several considerations, as listed below.

- *The effect of a pathogenic variant on gene function*. The simplest situation occurs when the variant affects how a single gene functions, which we consider in the section below. Some variants, however, simultaneously affect, directly or indirectly, how multiple genes work. A large-scale mutation, for

example, can directly change the sequence or copy number of multiple neighboring genes. In addition, a simple mutation in a regulatory gene, such as an miRNA gene, can indirectly affect the expression of multiple target genes, and can potentially cause complex phenotypes.

- *The extent to which normal copies or different copies of the mutant gene are also available*. For diploid nuclear genes, that means considering how a mutant allele and a normal allele work in the presence of each other, or estimating the combined effect of two mutant alleles. The situation for mitochondrial DNA mutants is quite different, and rather unpredictable, partly because we have so many copies of mtDNA in each cell, and partly because, unlike nuclear DNA, mtDNA replication is not governed by the cell cycle.
- *The effect of other interacting factors*. Affected genes do not work in isolation: many factors—non-allelic genetic factors (other loci), epigenetic factors, and environmental factors—affect the extent to which an individual pathogenic variant affects the phenotype. The situation becomes even more complicated when multiple genetic variants at different loci are involved—we examine this in Chapter 8, within the context of complex disease.

Mutations affecting how a single gene works: an overview of loss of function and gain of function

Mutations that affect how a single gene works can have quite different effects on how the gene makes a product. Some affect expression levels only, often causing complete failure to express the normal gene product, or causing a substantial reduction in expression. Mutations that increase gene copy number and occasional activating point mutations result in overexpression (which can be a problem for *dosage-sensitive genes*). Other mutations can result in an altered gene product that lacks the normal function, or that has an altered or new function.

One broad way of classifying the overall effect of a mutation is to consider whether the mutation results in a loss of function or a gain of function, as described below. As described in the section after this one, the effect of a loss-of-function mutation is often minimal in the presence of a normal allele, but a gain-of-function mutation has a harmful effect even in the presence of a normal allele.

Loss-of-function mutations

Loss-of-function mutations can have different consequences for how a gene is expressed. Sometimes, the final gene product is simply not produced or is nonfunctional (in which cases, the mutant gene is said to be a **null allele**). Different types of mutation can produce null alleles, notably large-scale deletions (eliminating the entire gene or a significant portion of it). Various types of mutation in protein-coding genes introduce early premature termination codons, causing the mRNA to be degraded so that no protein is made (as detailed in Box 7.1). They include many nonsense mutations, and various frameshifting mutations operating at either the DNA level (insertions or deletions), or at the RNA level, either via exon skipping or intron retention (as shown in Figure 7.3), or via exon truncation or exon extension (Figure 7.4).

For some loss-of-function mutations, there is some residual activity (the gene is expressed at abnormally low levels because of mutations in a regulatory sequence), or it is expressed normally but works poorly (as a result of a small non-frameshifting insertion, for example). A small minority of loss-of-function mutations are missense mutations that replace a key amino acid by a rather different amino acid. Recall that specific amino acids can have critical functional roles (including in post-translational processing), and structural roles (certain cysteines participate in disulphide bonding, for example).

Gain-of-function mutations

Gain-of-function mutations typically give rise to products that are positively harmful in some way. They are common in cancers, but in inherited disorders they are usually much less common than loss-of-function mutations. An outstanding exception to this general rule is provided by the paternal age effect disorders previously encountered in Table 7.4. Each of these disorders is caused by missense mutations that activate a member of the growth factor receptor–RAS signal transduction pathway so as to confer a selective growth advantage on spermatogonial stem cells. The end result is clonal expansion of spermatogonial stem cells containing the mutant gene—see **Figure 7.18** for the example of mutations in the *FGFR3* (fibroblast growth factor receptor 3) gene, causing bone dysplasia.

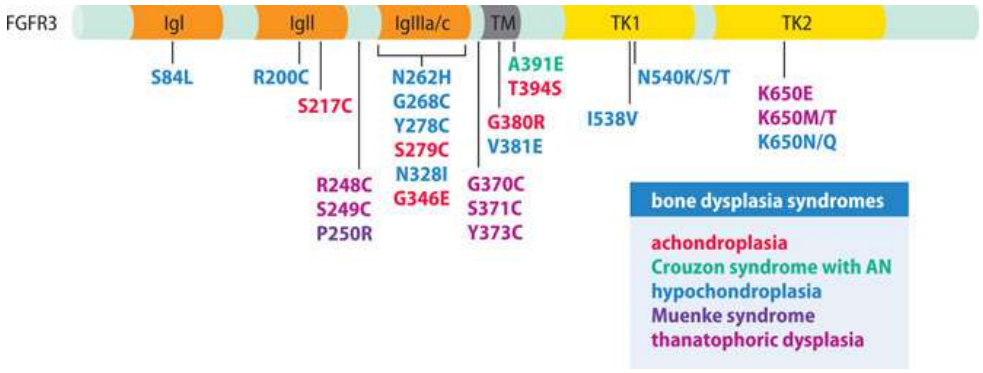


Figure 7.18 Pathogenic amino acid substitutions in the FGFR3 protein due to gain-of-function missense mutations that confer a selective growth advantage on spermatogonial stem cells. The fibroblast growth factor receptor 3 (FGFR3) protein has three immunoglobulin-like domains (Igl, Igll, and Iglla/c) in the extracellular region (shown on the left), a single hydrophobic transmembrane (TM) domain, and two cytoplasmic tyrosine kinase domains (TK1, TK2). The indicated amino acid substitutions shown below cause a type of bone dysplasia syndrome, according to the color scheme in the box. They arise from gain-of-function mutations that appear to confer a selective advantage on spermatogonial stem cells containing the mutant allele. AN, acanthosis nigricans. (Adapted from Goriely A & Wilkie AO [2012] *Am J Hum Genet* 90:175–200. With permission from Elsevier.)

For inherited monogenic disorders other than the examples showing paternal age effect, gain-of-function mutations can work in different ways. Some produce a radically altered product. Recall the examples in Section 7.3 where unstable dynamic mutations give rise to toxic proteins (which kill cells), or harmful *trans*-dominant RNAs (which cause disease after inappropriately binding to proteins made by other genes).

Most gain-of-function mutations do not produce a radically new product. Instead, they tend to make products that are structurally abnormal (and cause aberrant protein folding or aggregation), or are expressed inappropriately in some way. For example, a cell-surface receptor might be inappropriately expressed in the wrong cells, so that a particular signaling pathway is available when it should not be. As a result, the ectopically or inappropriately expressed product is able to interact with other cell components that it normally might not interact with, causing disease.

Sometimes it is a question of mutation causing altered specificity in some protein. A good example is the Pittsburgh variant of α1-antitrypsin (a1-AT), a member of the serpin family of serine protease inhibitors, which cleave its target proteins, potentially dangerous serine proteases, at specific sites. An important function of α1-antitrypsin (a1-AT) is to protect normal tissues from high levels of elastase, a

protease expressed by neutrophils during inflamma-tory responses (high elastase levels trigger a compensatory increased α1-AT production to suppress elastase). Elastase cleaves a specific peptide bond in the α1-AT backbone that activates a radical conformational change in the a1-AT molecule, leading to elastase destruction. In the Pittsburgh variant of α1-AT a missense mutation changes the specificity of the enzyme, and the mutant α1-AT now attacks a different serine protease, the blood clotting factor thrombin, causing a lethal bleeding disorder (interested readers can find a detailed account at PMID 11778003).

Gain-of-function mutations are particularly common in cancer; many of these arise from chromosomal translocations and other rearrangements that create chimeric genes. Chromosomal translocations cause major problems in meiosis and so they are rarely responsible for inherited disease. But cancers arise from a mitotic division of mutant somatic cells in which the chromosomal rearrangements can be readily propagated from mother cell to daughter cells.

The effect of pathogenic variants depends on how the products of alleles interact: dominance and recessiveness revisited

Most of the genes in our diploid cells are present in two copies—one inherited from the mother and one from the father. For heterozygotes, therefore, we need to consider how the expression of a pathogenic variant (mutant allele) might affect the normal allele. To what extent might the effect of the mutant allele be reduced by—or compensated for by—having a normal allele? And, secondly, can a mutant allele have an adverse effect on how the normal allele is expressed?

Loss-of-function versus gain-of-function mutations in recessive and dominant disorders

Recall that in dominant conditions the disease phenotype is somehow expressed in the heterozygote, but in recessive conditions heterozygotes are unaffected. In autosomal recessive conditions, therefore, the disorder is expressed only when both alleles are pathogenic variants. Heterozygous carriers cannot have a gain-of-function mutation (otherwise they would be expected to be affected) and so have one loss-of-function mutation, and affected individuals have two alleles with loss-of-function mutations.

Dominant conditions are less uniform. There is one mutant allele and, according to the disorder, the mutant allele may be a gain-of-function or a loss-of-function mutation. It is easy to imagine how a gain-of-function mutation might work in a heterozygote if we think of it as being positively harmful, causing damage even if the other allele is pumping out the correct product. Think of the toxic products produced by many unstable short tandem repeat expansions—the presence of functional product made by the normal allele is not going to stop the toxic effects of the mutant allele. (Potential gene therapies that rely on providing normal alleles could never work; instead we would have to inhibit the harmful effects of the mutant allele in some way.)

But how does just one loss-of-function mutation cause a dominant disorder? Why is the normal product made by the unaffected allele not enough? In a few cases, such as for imprinted loci (Section 6.3), only one of the two alleles is normally expressed, and the unaffected allele happens to be the one that is silenced. Thus, for example, a single loss-of-function mutation of the maternal *UBE3A* allele is enough to cause Angelman syndrome—the paternal *UBE3A* allele is not expressed (at least, not in the brain).

In most cases, both alleles of a diploid gene locus are normally expressed, but a single loss-of-function mutation can nevertheless cause disease if the gene concerned shows exceptional dosage sensitivity (**Box 7.3**). (Note that for all dominantly inherited disorders caused by a loss-of-function mutation, the phenotype is recessive at the cellular level: the normal phenotype can be restored by introducing a functioning allele or by reactivating a silenced allele.)

BOX 7.3 DOSAGE-SENSITIVE GENES AND HAPLOINSUFFICIENCY

A small number of our genes are not essential (people with blood group O, for example, have two inactive alleles at the *ABO* gene locus without any harmful effects). For very many single-copy genes, however, homozygous inactivation is a problem, and complete absence of a gene product often results in disease or is lethal. The amount of product made by most genes can, however, show significant variation without harmful effects, if above some critically low level.

For a diploid locus, an occasional gene duplication—resulting in a total of three gene copies and an expected 50 % increase in gene product—often makes no obvious difference to the phenotype. According to circumstances, overproduction

can even be advantageous: Western populations that traditionally have diets rich in starch have many copies of the starch-processing a-amylase gene, *AMYIA* (see Figure 4.8).

Similarly, for many genes, a reduction to 50 % of gene product is often inconsequential. In recessive loss of function, one null allele at a diploid locus typically causes no harm (provided the other allele works normally). Even if the second allele also has a hypomorphic mutation leaving it only partly functional, and the combined output of the two alleles is, say, 30 % of the normal amount of gene product, there may be little evidence of pathogenesis (**Figure 1**). For these genes, the product needs to drop to rather low levels before disease becomes apparent. (This especially applies to products such as enzymes that can be used over and over again.)

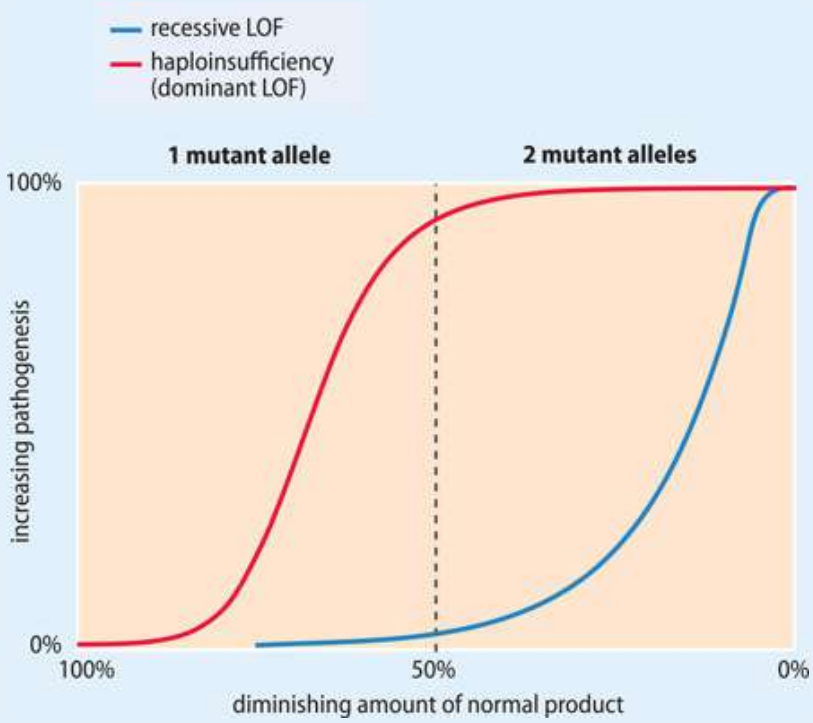


Figure 1 The relationship of disease susceptibility to diminishing amounts of gene product in dominant and recessive disorders due to loss-of-function mutations. The vertical dotted line marks the point at which an individual has a fully functional normal allele and a null allele (such as a gene deletion) and might be expected to make about 50 % of the normal amount of gene product. If the allele is not null but partly functional, the amount of product made increases (as shown on left). If the second allele also has a loss-of-function (LOF) mutation, the amount of gene product made diminishes, depending on the severity of that mutation (as shown on the right). For dosage-sensitive genes, the reduction from 100 % to 50 % gene product is sufficient to result

in disease (haploinsufficiency). Increasing the amount of gene product (by having three gene copies) can also induce disease. In recessive disorders, the disease is normally manifested only when both alleles have loss-of-function mutations; pathogenesis increases rapidly as the amount of normal product made approaches zero. There is some variability in recessive diseases—for some disorders, heterozygotes may show limited pathogenesis. Note that the idealized curves shown here do not take into account other factors such as modifier genes or environmental factors.

A minority of our genes are especially **dosage-sensitive**—the amount of product made is critically important. Changes in gene copy number (gene *dos- age*) can cause disease by changing the amount of gene product beyond normal limits, and certain types of point mutation can have the same effect by reducing or amplifying gene expression. Disease can occur when too much of a product is made; sometimes that happens when the increase is only 50 % above the normal amount. For example, as described below, type 1A Charcot–Marie–Tooth disease (a hereditary motor and sensory neuropathy) is often caused by a duplication giving rise to three copies of the *PMP22* gene, or by point mutations that lead to the overexpression of *PMP22*.

More commonly, disease is due to a loss-of-function mutation in one allele of a dosage-sensitive gene (the other allele can be normal and expressed). If the mutant allele is a null, such as a gene deletion, the amount of normal gene product might be expected to be reduced by about 50 % (but pathogenesis can sometimes be observed when the mutant allele retains partial function). Because heterozygotes are affected, this type of loss of function is dominantly inherited, and is known as **haploinsufficiency** (see Figure 1 in Box 7.3). Heterozygotes for a loss-of-function mutation in a dosage-sensitive gene are rare, and so having two loss-of-function alleles—usually by being a compound heterozygote—is extremely rare. When observed, the phenotype is often slightly more severe than that for a heterozygote.

Dosage-sensitive genes typically make products that need to be calibrated against the level of some other interacting or competing gene product. In many cases, the products have roles in quantitative signaling systems or other situations in which precisely defined ratios of the products of different genes are important for them to work together effectively. Genes that regulate other genes are likely candidates: they might do so by making transcription factors, signaling receptors, splicing regulators, or chromatin modifiers, for example. Or the different gene products may be

antagonistic, competing with each other to ensure that some critical reaction is carried out that is important in development or metabolism. Because chromosomes usually have multiple dosage-sensitive genes, constitutional aneuploidies are often lethal, but some are viable (as shown in Table 7.12).

Striking loss of function produced by dominant-negative effects in heterozygotes

In heterozygotes a null allele causing a loss of function does not normally affect the function of the normal allele. Sometimes, however, a mutation results in a mutant protein that cannot perform the normal function and also inhibits the function of protein produced by the normal allele. A mutant allele making a protein that antagonizes the protein made by the normal allele in a heterozygote is sometimes known as an *antimorphic* allele, and provides an example of a **dominant-negative** effect.

A common example occurs when the normal protein is part of a multimer that is inactive if it incorporates any of the mutant protein. Imagine the simplest possible case, when the multimer is a homodimer. A heterozygote for a null allele might be expected to make 50 % of the normal homodimer. However, a heterozygote who makes equal quantities of the normal monomer and a mutant monomer (which can only form inactive dimers) might be expected to make only 25 % of the normal amount of functional dimer (**Figure 7.19A**).

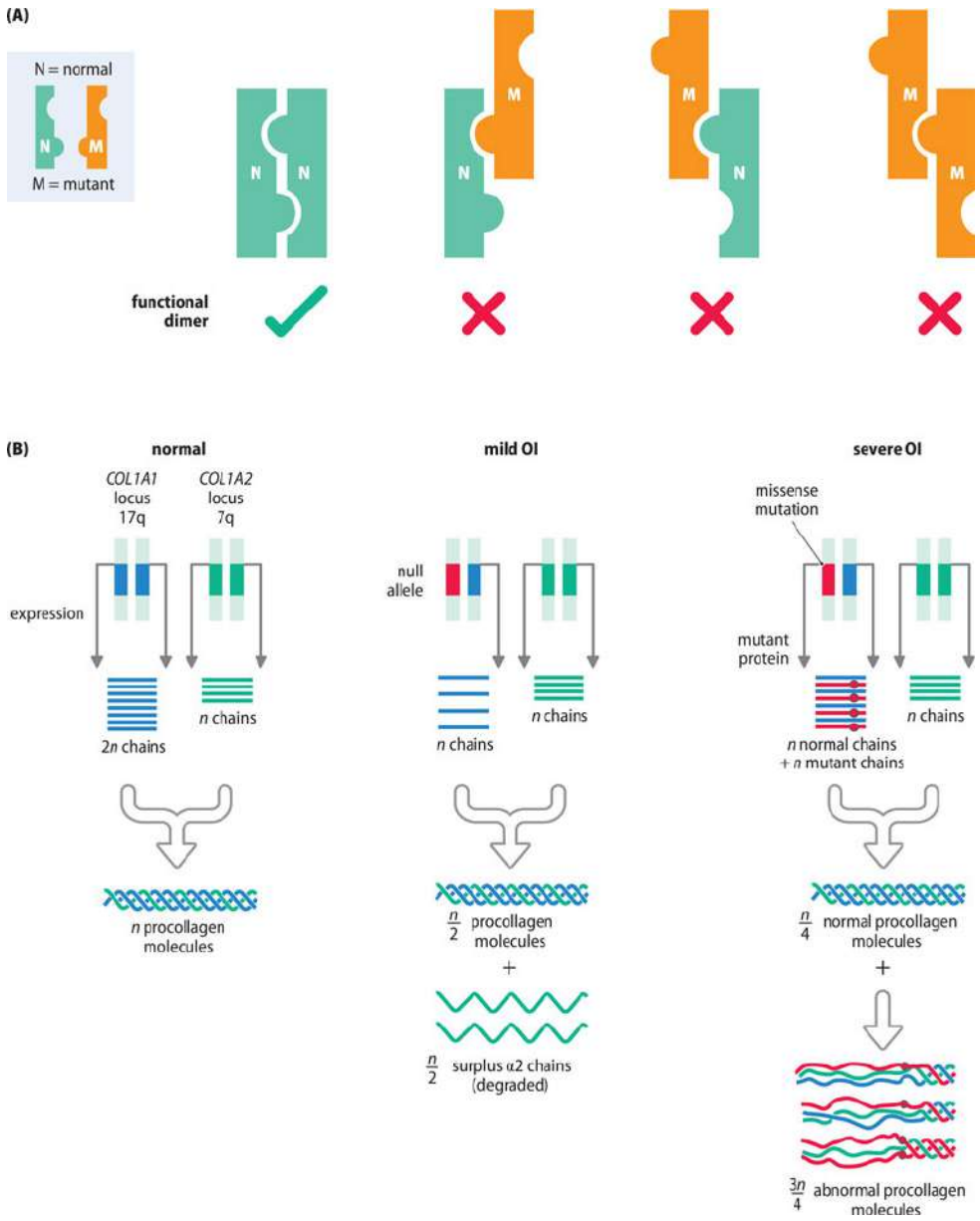


Figure 7.19 Dominant-negative effects: when a heterozygous missense mutation is more harmful than a null allele. (A) A hypothetical example: the product of a gene forms a homodimer, and the mutated allele produces a protein at normal quantities that can only form nonfunctional dimers. As a result, only one-quarter of the normal amount of functional dimers is made (there are two ways of forming nonfunctional heterodimers and one type of nonfunctional mutant homodimer). (B) A clinical example: a dominant-negative mutation causing a severe type of osteogenesis imperfecta (OI). Two *COL1A1*-encoded polypeptides and one *COL1A2-encoded* polypeptide are required to make a triple-helical type I procollagen. A null allele at *COL1A1* simply reduces the amount of type I procollagen by half and results in a mild form of OI. However, mutations that replace a structurally important glycine with any other amino acid usually have

strong dominant-negative effects because they disrupt the packing of the three chains in the triple helix. The mutant *COL1Al-encoded* polypeptide is included within three-quarters of the type I procollagen molecules made, making them nonfunctional. With only 25 % of the normal type I procollagens, affected individuals develop a severe type of OI.

Disorders of structural proteins that work as multimers often provide disease phenotypes arising from dominant-negative proteins, including examples in osteogenesis imperfecta (collagens), Marfan syndrome (fibrillins), and epidermolysis bullosa (keratins). As an illustration, consider collagens in osteogenesis imperfecta (also called brittle bone disease). Collagens are initially synthesized as procollagens in which three polypeptide chains are wound round each other, beginning at the C-terminal end, to form a stiff, rope-like triple helix. Each helical polypeptide chain undergoes a remarkable one turn every three amino acids, which is made possible because collagen polypeptides have a unique structure based on tandem repeats of a three-residue sequence with the general formula Gly–X–Y. The glycines, with a single hydrogen atom in the side chain, provide the flexibility; X and Y are often (but not always) proline and hydroxyproline, respectively, which have unique side chains that loop back and connect to the polypeptide backbone, stabilizing the helical structure.

In type I procollagen, the precursor of the most common type of collagen, two of the three polypeptide chains are made by the *COL1A1* gene and the third is made by another gene, *COL1A2*. A null mutation in *COL1AI* (or *COL1A2*) results in a mild form of osteogeneis imperfecta—in each case the amount of procollagen made might be expected to decrease to 50 % (Figure 7.19B). By contrast, a mutation that replaces a glycine in a collagen polypeptide by any other amino acid usually has dominant-negative effects: it causes abnormal packing of the collagen polypeptides into a triple helix for any collagen molecule into which it is incorporated. Such a mutation in the *COL1A1* gene causes a severe form of osteogenesis imperfecta, the type IIA form, because the amount of procollagen produced would be expected to decrease to just 25 % (see Figure 7.19B).

The antagonistic effect of an antimorph in preventing wild-type protein from performing its function might be considered a gain of function, but as shown in Figure 7.19B the net effect is to create a greater loss of function than a null allele.

Gain-of-function and loss-of-function mutations in the same gene can produce different phenotypes

Phenotypes arising from a loss of function are typically associated with mutational heterogeneity (there are many different ways of inactivating a gene—frameshifting mutations, nonsense mutations, major splice-site mutations, missense mutations, and whole gene deletions). Gain-of-function mutations are less common in inherited disorders, and the resulting phenotypes are typically associated with mutational homogeneity. They may be a class of unstable oligo-nucleotide expansions, for example, or specific activating missense mutations, or a mutation that results in overexpression, but not a great range of different types of mutation.

Gain-of-function mutations and loss-of-function mutations in the same gene quite often produce very different phenotypes. In some cases, an inherited disorder is caused by the loss-of-function mutation, but a gain-of-function mutation in the same gene produces a cancer. For example, loss-of-function mutations in the *RET* gene (which makes a tyrosine kinase) result in susceptibility to Hirschsprung's disease in which affected individuals have a congenital absence of ganglia in the gut. But different, very specific, kinds of activating missense mutations in the same gene result in different types of cancer: either medullary thyroid carcinoma or multiple endocrine neoplasia types 2A or 2B.

Loss-of-function mutations in the androgen receptor gene (*AR*) cause androgen-insensitivity syndrome (also called testicular feminization syndrome). Affected individuals have a 46,XY karyotype but because their androgen receptors do not work normally the end organs are insensitive to androgens, resulting in an X-linked recessive form of pseudohermaphroditism. Exon 1 of the *AR* gene also happens to have a tandem CAG repeat that can undergo unstable expansion to produce androgen receptor proteins with an expanded polyglutamine tract. The resulting proteins (and possibly RNA transcripts) are toxic to vulnerable cells and lead to spinal and bulbar muscular atrophy (also called Kennedy disease). In this case there is degeneration of lower motor neurons affecting certain muscles in the arms and legs and also some muscles in the face and throat (bulbar muscles).

Another case of divergent gain-of-function and loss-of-function phenotypes in one gene is provided by the *PMP22* gene, which makes a peroxisomal membrane protein. Because the great majority of pathogenic mutations are duplications or deletions of a 1.4 Mb region at 17p11.2 that contains multiple genes in addition to *PMP22,* we consider the pathogenesis in the next section.

Multiple gene dysregulation resulting from aneuploidies and mutations in regulatory genes

Some genes produce regulatory proteins or RNAs that regulate many different target genes. Examples include genes encoding high-level gene regulators (making master transcription factors, microRNAs or splicing regulators) and genes involved in global epigenetic regulation, such as those that make chromatin modelers or DNA methyltransferases. These genes are typically dosage-sensitive, and heterozygotes with a loss-of-function mutation can often show complex phenotypes (see Table 6.7 on page 167 for some examples).

Whole-chromosome aneuploidies and large-scale intrachromosomal deletions and duplications (*segmental aneuploidies*) also directly affect multiple genes simultaneously, in this case by changing their copy number. However, because most genes are comparatively insensitive to dosage effects, the pheno-type in affected heterozygotes is due to the combined effects of a comparatively small number of dosage-sensitive genes.

Whole chromosome aneuploidies

Because the number of dosage-sensitive genes over a whole chromosome is often large, several dosage-sensitive genes might be present; reducing the gene copy number by 50 % could be expected to have severe consequences. Monosomies are almost always lethal because the cumulative effect of deleting one copy of each dosage-sensitive gene on a chromosome is simply too much to support embryonic or fetal development.

One monosomy can sometimes be viable. The 45,X genotype leads to spontaneous abortion in about 99 % of cases, but occasionally leads to Turner syndrome, a comparatively mild condition: short stature with certain minor physical abnormalities (webbed necks, low-set ears) and gonadal dysfunction causing sterility. X-chromosome inactivation means that 45,X women are normally functionally monoallelic for most X-linked genes, but a few genes on the X chromosome, including genes in the pseudoautosomal regions, are not subject to X-inactivation, however.

Providing an extra gene copy for dosage-sensitive genes might be expected to have less harmful effects, but across a whole chromosome the combined effects of dosage imbalance in multiple genes means that most autosomal trisomies are also lethal. Chromosomes with few genes have fewer dosage-sensitive genes, and the three autosomal trisomies compatible with life—trisomies 13, 18, and 21—each involve chromosomes with relatively few genes.

Having extra sex chromosomes has far fewer ill effects than having an extra autosome because of X-inactivation (inactivating all X chromosomes except one) and the scarcity of genes on the Y chromosome. People with 47,XXX and 47,XYY karyotypes often function within the normal range, and in comparison with people with an autosomal trisomy, men with 47,XXY (Klinefelter syndrome) have relatively minor problems, notably hypogonadism and reduced fertility.

Segmental aneuploidies

Large-scale subchromosomal deletions and duplications also cause disease by simultaneously changing the copy number of multiple linked genes. If they occur in chromosomal regions that are constitutionally hemizygous, or have genes showing monoallelic expression (via imprinting, X-inactivation and so on), the functional copy number of some genes will be reduced to zero, so that no gene product is made at all. A profound effect can therefore often be expected for a few of our gene loci, however, a complete absence of gene product does not result in a clinical phenotype (people with blood group O, for example, have two inactive *ABO* alleles).

Males are constitutionally hemizygous for X- and Y-specific regions; large deletions in these regions therefore result in a complete absence of multiple gene products. The Y chromosome has few genes and they are very largely devoted to male-specific functions; accordingly, large deletions here are associated with azoospermia and infertility. Large deletions within the X chromosome in males are often lethal because of the high density of genes that perform a wide range of important functions. Certain regions such as Xp21 are comparatively gene-poor, however, and large deletions here can result in disease phenotypes.

Large-scale deletions and duplications on autosomes can cause disease by changing the copy number of comparatively rare dosage-sensitive genes. Deletions reduce the functional gene copy number to one (so that disease results from haploinsufficiency in dosage-sensitive genes), and duplications increase the gene copy number to three, resulting in the overexpression of multiple genes.

7.8 A PROTEIN STRUCTURE PERSPECTIVE OF MOLECULAR PATHOLOGY

Until now we have looked at pathogenesis mostly from the perspective of altered gene expression or changes in protein sequence that either result in loss of function or produce a gain of function that does not necessarily involve a major change in protein structure. However, for many disorders the pathogenesis is ultimately due to major changes in protein *structure* that are typically induced by single nucleotide substitutions or other point mutations.

We briefly described elements of protein structure and folding in Chapter 2. In the next two sections we consider disease caused when proteins adopt altered structures and when changes in structure can predispose proteins to form aggregates that can cause disease. Understanding the basis of major changes in protein structure is important for understanding molecular pathology and developing novel therapeutic strategies.

Pathogenesis arising from protein misfolding

We have previously considered various aspects of how the expression of protein-coding genes is regulated to give a functional product, dwelling mostly on transcriptional, post-transcriptional, and translational control in Chapter 6, and touching on mRNA surveillance in Box 7.1. However, for proteins to function correctly they also need to fold properly to assume the correct three-dimensional conformation so that they can bind the appropriate interacting molecules. They need to function correctly within the right environment (in a hydrophilic environment, proteins fold up with hydrophobic amino acids located in the interior, and hydrophilic amino acids on the surface). Proteins also need to be able to interact correctly with other proteins when forming multimers.

Regulation of protein folding

Protein folding is not straightforward: there are various different paths that can be taken by an unfolded or partly folded protein to arrive at the final conformation, and natural errors in protein folding are common. Some proteins can fold correctly without help, but many proteins require assistance in folding from specific molecular chaperones such as Hsp60 or Hsp70 (chaperones are sometimes called *heat-shock proteins*, with their names prefixed by *Hsp,* because their expression is drastically increased when cells are exposed to even moderate temperature increases, such as from 37°C to 42°C, that nevertheless cause an increase in protein misfolding).

Chaperones can help to fold both incompletely folded proteins and some incorrectly folded proteins. However, when attempts to refold a protein are unsuccessful, the protein is shunted into a proteolytic pathway in which it is destroyed by the *proteasome*, a complex compartmentalized protease that is distributed in many copies throughout the cytosol.

Aberrant protein folding causing disease

Protein misfolding is a common cause of disease in many genetic disorders, such as cystic fibrosis and phenylketonuria, in which mutations that change a single amino acid are quite common. Thus, about 90 % of individuals with cystic fibrosis have one or two copies of the p.Phe508del allele in which deletion of a single phenylalanine residue is sufficient to cause aberrant protein folding that cannot be rectified by chaperones. Whereas the normal protein would continue its regular journey to take up residence in the plasma membrane, the mutant protein is rapidly subjected to degradation in the endoplasmic-reticulum.

Sometimes mutations destroy the ability to adopt highly specific structures required for assembly of multimers such as collagens, fibrillins, and keratins. Collagens, for example, require complex packaging of three collagen polypep-tides into trihelical structures in which three individual collagen strands are wound round each other. As described above, missense mutations that replace glycines in collagen chains cause major packaging problems (as shown previously in Figure 7.19B).

The many different ways in which protein aggregation can result in disease

The pathogenesis of many disorders, both monogenic and common diseases, involves protein aggregation. Soluble oligomers and large, often insoluble complexes can form and the aggregated proteins are often found as cellular inclusions or pericellular deposits.

At present there is still some uncertainty about the significance of protein aggregates observed in many common diseases—are they a direct cause of disease, or are they more peripherally associated with pathogenesis? In monogenic disorders there can be more certainty, although sometimes even here the precise pathogenetic process is still not clear. In some monogenic disorders, however, the evidence for mutation-induced protein aggregation is clear; we give two examples below, one in

which proteins aggregate to form extremely long protein fibers, and one in which the damage is done by protein aggregates in inclusion bodies within cells.

We end this section with what used to be thought of as a bizarre protein aggregation disease mechanism, one involving a type of epigenetic information transfer *that does not involve nucleic acids* (and so is quite distinct from the epigenetic mechanisms described in Section 6.3, which rely on heritable chromatin states). The first evidence came from studies of mutant prion proteins, but as described below similar pathogenetic mechanisms are now thought to occur in various disorders including some common neurodegenerative diseases such as Alzheimer and Parkinson disease.

Sickle-cell anemia: disruptive protein fibers

Normal adult hemoglobin is a tetramer with two a-globin chains and two b-globin chains. Individuals affected by sickle-cell anemia are homozygous for a specific missense mutation that replaces a charged, hydrophilic glutamate residue at position 6 in the b-globin chain by a hydrophobic valine residue. The resulting mutant hemoglobin S (HbS) has a strong tendency to aggregate when deoxygenated, resulting in fibers composed of 14 long strands of HbS tetramers (**Figure 7.20A,B**). The fibers cause red blood cells to be deformed so that they are crescent-shaped, like a sickle. The much shorter life span of the abnormal sickle cells (10–20 days, in contrast with the normal 90–120 days) means that the body cannot replace dead red blood cells fast enough, resulting in anemia. The HbS fibers also block small blood vessels, causing hypoxic tissue damage.

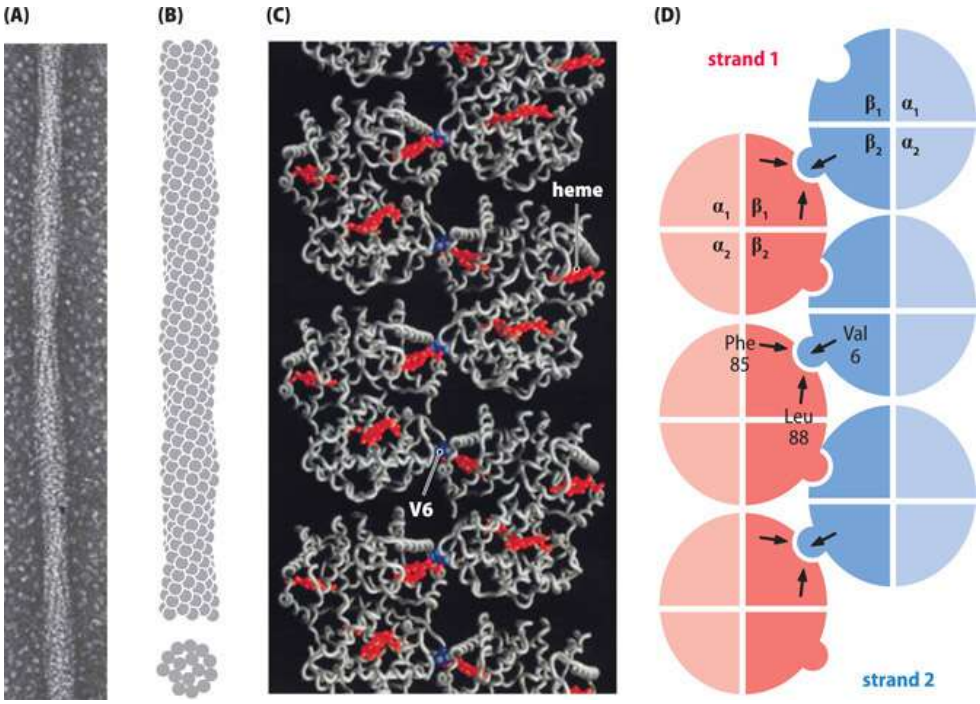


Figure 7.20 Aggregation of hemoglobin to form complex fibers in sickle-cell disease. (A,B) The 14-strand structure of deoxyhemoglobin S fibers including an electron micrograph in (A) of a stained fiber and the interpreted structure in (B), showing a lateral image at the top and a cross section at the bottom. The 14-strand structure is built from seven sets of paired strands. (C,D) The basic paired strand component of deoxyhemoglobin S fibers. (C) Structural model showing the mutant valine (V6; in blue color) located on the outside of the b-globin chains, facilitating lateral contact between b-globin chains on different HbS tetramers. Heme groups are shown in red. (D) Diagram illustrating how each double strand of hemoglobin tetramers is stabilized by lateral contacts involving the mutant valine on a b-globin chain from one strand interacting with a pocket formed between two helices on a b-globin chain on a Hb tetramer on the opposing strand. (A,B, From Dykes G et al. [1978] *Nature* 272:506–510; PMID 692655. With permission from Macmillan Publishers Ltd; C, from Harrington DJ et al. [1997] *J Mol Biol* 272:398–407; PMID 9325099. With permission from Elsevier.)

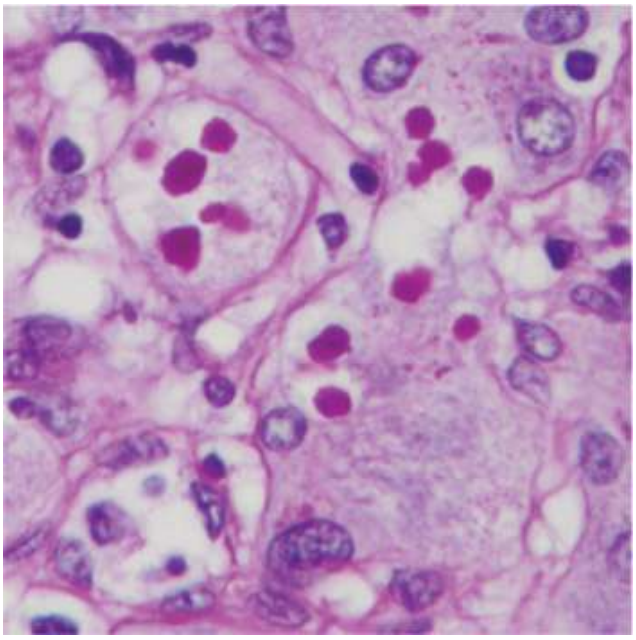
The 14-strand structure is built on the lateral association of seven sets of paired HbS tetramer strands. The side chain of mutant valines on the b-globin chains of one HbS strand can interact with a complementary pocket on b-globin residues of the neighboring HbS tetramer. That type of bonding drives the formation of paired strands of HbS tetramers (Figure 7.20C,D) that then form the higher-order structure shown in Figure 7.20A,B by additional lateral associations.

α1-Antitrypsin deficiency: inclusion bodies and cell death

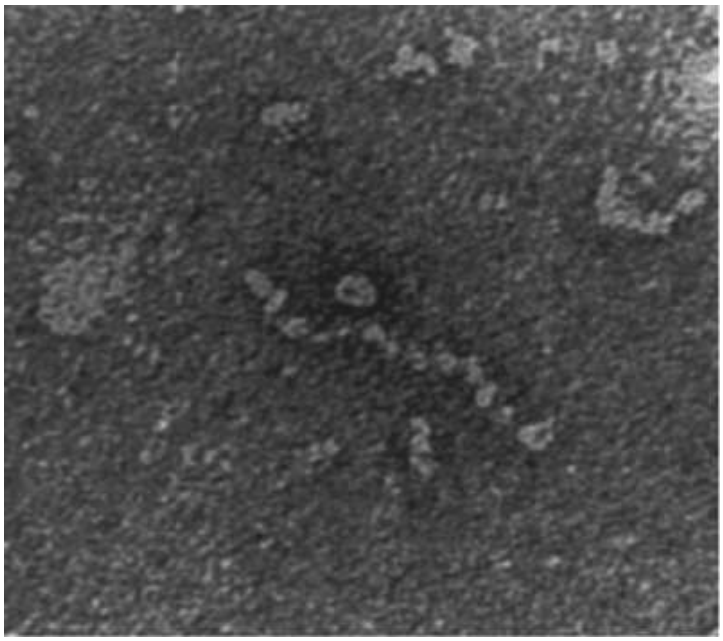
α1-Antitrypsin (α1-AT) is made by the liver and secreted to regulate the levels of certain serine proteases such as elastase (which can be overproduced by neutrophils during inflammation and might damage sensitive tissue, such as the alveoli of lungs, if not kept in check). α1-Antitrypsin deficiency is common in Caucasian populations in which two missense mutations are especially common: the mild PI*S allele (E264V) and the severe PI*Z allele (E342K).

Plasma concentrations of α1-AT in ZZ homozygotes (about 15 % of normal) and in SZ compound heterozygotes (about 40 %) are not high enough to protect lungs from damage by elastase over a lifetime, especially in people who smoke. Affected individuals often develop emphysema, a form of chronic obstructive lung disease in which tissues needed to support the shape and function of the lungs are destroyed. The low plasma α1-AT concentrations typically do not result from failure of liver cells to make any of the protein; instead the problem is a blockage in α1-AT processing and secretion from liver cells.

The retained α1-AT proteins can aggregate in the endoplasmic reticulum of hepatocytes to form intracellular inclusions (*inclusion bodies*) that can be readily recognized by using suitable stains and can be seen to contain bead-like polymerases in ZZ homozygotes (**Figure 7.21**). The inclusion bodies cause hepatocytes to die and can result in eventual cirrhosis of the liver, especially in ZZ homozygotes.



(A)



(B)

Figure 7.21 Intracellular inclusion bodies and protein aggregates in α1- antitrypsin deficiency. (A) Staining of hepatocytes with a periodic acid-Schiff stain reveals inclusion bodies as bright pink globules (arrowed). (B) Electron microscopy showing bead-like polymers of Z-type α1-antitrypsin. (A, Courtesy of the National Society for Histotechnology; B, from Lomas DA et al. [1993] *J Biol Chem* 268:15333–5; PMID 8340361. With permission from The American Society for Biochemistry and Molecular Biology.)

Seeding of mutant protein using aberrant protein templates

In Section 6.3 we explored epigenetic gene regulation in inherited disorders. That relies on heritable chromatin states rather than the DNA sequence, but yet another type of information that directs heritable changes in gene products is confined to the protein level. For some types of protein, mutant proteins that can aggregate are also able to direct normal forms of the same protein to adopt the mutant protein structure, allowing cellular spreading of protein aggregation. The first examples came from prion diseases in livestock, which garnered much public attention because of the danger to health in eating meat from infected cattle ("mad cow disease"). Similar mechanisms allow the cellular spreading of protein aggregation in certain other neurodegenerative diseases, including some monogenic disorders and also complex diseases including Alzheimer disease and Parkinson disease (see **Clinical Box 8**).

CLINICAL BOX 8 PRION DISEASES AND PRION-LIKE NEURODEGENERATIVE DISEASES: SEEDED SPREADING OF PROTEIN AGGREGATES BETWEEN ORGANISMS AND CELLS

Prion diseases—also known as transmissible spongiform encephalopathies—are progressive, fatal, and incurable neurodegenerative disorders that affect humans and other animals, in which holes develop in brain tissues, giving them a sponge-like texture. The disease can be spread from one organism to another by ingesting or internalizing affected tissue. For example, consumption of affected tissue from cows with bovine spongiform encephalopathy has led to outbreaks of variant Creutzfeldt–Jakob disease (vCJD; also known in this specific instance as "mad cow disease"). In addition to acquired prion protein disease, sporadic and hereditary forms exist. Creutzfeldt–Jakob disease (CJD), fatal familial insomnia, and Gerstmann–Straussler–Scheinker syndrome are dominantly inherited allelic disorders resulting from mutations in the *PRNP* prion protein gene at 20p13.

In prion disease, a normal cellular form of prion (PrPC) is misfolded into an abnormal conformation (PrPSc) that is rich in b-pleated sheets and prone to aggregation (Figure 1A; the Sc superscript comes from scrapie, a sheep prion disease that was one of the first to be studied).

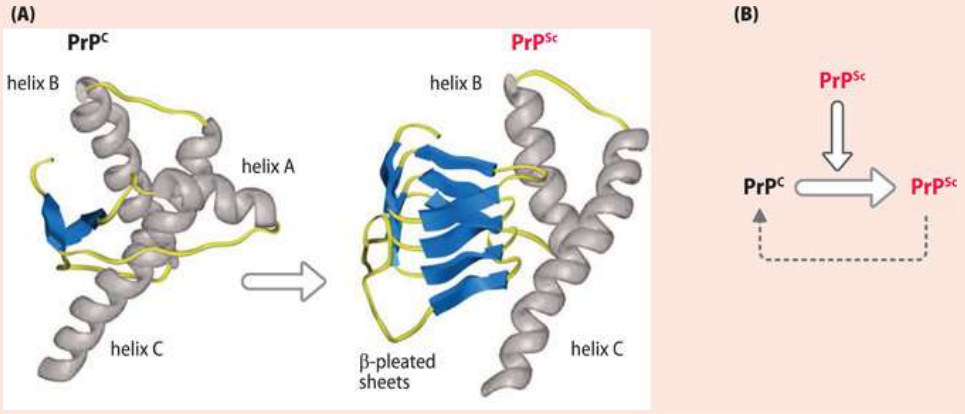


Figure 1 Conversion of normal prion protein (PrP^C) to the aggregation-prone isoform PrP^{Sc}. (A) The normal PrP^C protein has three a-helices (A, B, and C in gray) and two very short b-strands (blue), but it can misfold to give the PrP^{Sc} isoform; because it has a very high content of b-strands that form a b-pleated sheet, the PrP^{Sc} isoform is susceptible to aggregation. (B) A PrP^{Sc} isoform that may have been induced by mutation, arisen spontaneously, or be derived from another person or an animal with prion protein disease (for example after blood transfusion or the ingestion of affected animal tissue) can induce a normal human PrP^C protein to change to the

infectious PrP^{Sc} isoform. The latter can then induce other host PrP^{C} proteins to convert to PrP^{Sc}, spreading the disease between cells. (A, Adapted from Norrby E [2011] *J Intern Med* 270:1–14. With permission from John Wiley and Sons, Inc.)

The most striking characteristic of PrPSc is that when it comes into contact with normal PrPC proteins it can induce them to switch conformation so that they, too, adopt the PrPSc structure. Thus, if our cells are exposed to abnormal prion proteins from an infected animal or person, the abnormal foreign prion proteins will induce host PrP^{C} proteins to adopt the PrPSc structure (Figure 1B).

The abnormal prion protein structure can effectively self-propagate by a form of replication that has nothing to do with nucleic acid sequences. In that respect the disease mechanism resembles classical epigenetic mechanisms (which typically involve chromatin modifications). The abnormal prion proteins are infectious because the misfolded protein can be acquired (by the ingestion of infected cells or tissue). Alternatively, prion proteins originate by a chance misfolding of a newly synthesized PrP^{C} protein in a sporadic case or develop as a result of a genetic mutation (in which the mutant sequence has a greater propensity to misfold).

The brain is the main target of prion toxicity—neurons, being extremely long-lived and not effectively replaced, are especially vulnerable to toxic protein aggregates. How prions enter the body and infect brain cells is an interesting question. Somehow, the abnormal protein aggregates can get past mucosal barriers, survive the attentions of innate and adaptive immunity, pass across the blood–brain barrier, and spread to different brain cells. Infection can be efficient. vCJD has been transmitted to hemophiliacs who were treated with a factor VIII extract isolated from blood samples provided by donors who included subclinically infected blood donors. Growth hormone deficiency and infertility have also been treated in the past with growth hormones or fertility hormones recovered from human cadaveric pituitary glands, but because the pituitary extracts had been contaminated by infected human brain tissue, more than 160 treated people died of vCJD.

AMYLOID DISEASES AND PRION-LIKE NEURODEGENERATIVE DISEASE

Prion proteins are members of the family of amyloid proteins that have a high content of b-sheets, making them prone to aggregation and the formation of elongated, unbranched amyloid fibrils. The spines of the fibrils consist of many-stranded b-sheets, arranged in a cross-b structure (**Figure 2**).

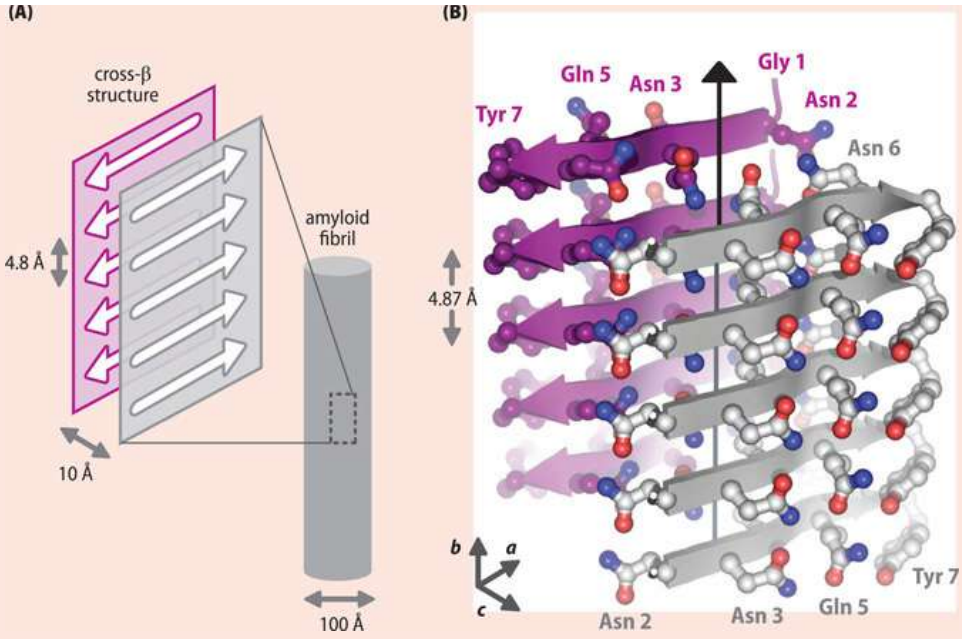


Figure 2 Characteristics of amyloid fibers. (A) Amyloid fibrils generally have a diameter of about 100 Å (10 nm) and form by protein aggregation. A characteristic property is the cross-b spine, a set of b-strands that run perpendicular to the axis of the fibril. (B) Certain protein segments that are six to seven amino acids long bind to other copies with the same amino acid sequence to form two tightly interdigitating b-sheets. In the case of the prion protein, that sequence is GNNQQNY or Gly-Asn-Asn-Gln-Gln-Asn-Tyr, as shown here. (Adapted from Eisenberg D & Jucker M [2012] *Cell* 148:1188–1203. With permission from Elsevier.)

Amyloid proteins are frequently associated with disease (Table 1), and the aggregates may be extracellular (PrP^{Sc}; b-amyloid), nuclear (huntingtin), or cytoplasmic (SOD1; Tau; and synuclein, which forms *Lewy bodies*). Amyloid protein aggregation is seen in some common diseases not associated with neurodegeneration, such as type 2 diabetes (where aggregates of serum amyloid A protein are found in the pancreatic islets of Langerhans. However, neurodegeneration is the most striking clinical characteristic of many amyloid diseases (Table 1).

TABLE 1 EXAMPLES OF AMYLOID DISEASES

Disease	Amyloid protein (precursor)

IAPP, islet amyloid polypeptide; APP, amyloid precursor protein; SOD1, superoxide dismutase 1.

* The non-amyloid protein TDP-43 is also commonly found to be aggregated in frontotemporal lobar degeneration.

Disease	Amyloid protein (precursor)
NON-NEURODEGENERATIVE DISORDERS	
Atherosclerosis	apolipoprotein Al
Rheumatoid arthritis	IAPP/amylin
Type 2 diabetes	serum amyloid A
NEURODEGENERATIVE DISORDERS	
Alzheimer disease	β-amyloid/Aβ (APP);Tau
Amyotrophic lateral sclerosis (motor neuron disease)	SOD1
Frontotemporal lobar degeneration (FTLD)-tau*	tau
Huntington disease	huntingtin
Parkinson disease	α-synuclein
Prion protein diseases	PrP^{Sc}(PrPC)

IAPP, islet amyloid polypeptide; APP, amyloid precursor protein; SOD1, superoxide dismutase 1.

* The non-amyloid protein TDP-43 is also commonly found to be aggregated in frontotemporal lobar degeneration.

Neurodegenerative amyloid diseases, such as Alzheimer disease, Parkinson disease, amyotrophic lateral sclerosis, and frontotemporal disease, resemble prion protein diseases in many ways and are sometimes classified as prionoid diseases. The direct involvement of the aggregated proteins in disease is supported from familial forms of these disorders in which mutations in the relevant gene promote the formation of amyloid protein, including mutations in *APP* (Alzheimer disease), *SNCA*, encoding a-synuclein (Parkinson disease), *SOD1* (amyotrophic lateral sclerosis), and *MAPT*, the microtubule-associated protein tau (frontotemporal dementia), for example.

There is no evidence from animal studies that the aggregated proteins in these disorders are infectious like prion proteins. But there is quite strong evidence that the pathogenesis resembles prion protein disease in two respects. First, like prion proteins, misfolded amyloid proteins in these disorders can induce the formation of the amyloid state in the normal proteins so that they aggregate. Secondly, for several of the disorders there is quite strong evidence for cell-to-cell spreading of the disorder.

7.9 GENOTYPE–PHENOTYPE CORRELATIONS AND WHY MONOGENIC DISORDERS ARE OFTEN NOT SIMPLE

Assessing the effect of pathogenic variants on the phenotype is a component of the broader quest to understand genotype–phenotype correlations. If we know the genotype, to what extent can we predict the phenotype? This can be a difficult question to answer, even for monogenic disorders.

The effect of a pathogenic variant on the phenotype is not just dependent on the effect of the mutation on the ability of that allele to make its normal gene product and the interaction with the product made by the other allele. Genes from other loci can also influence the disease phenotype, as can environmental factors. It is now clear that monogenic disorders are often not as simple as sometimes described, and the division of genetic disorders into chromosomal, monogenic, and multifactorial disorders is a simplification.

The difficulty in getting reliable genotype–phenotype correlations

Interpreting the effect of a single pathogenic mutation, even in a well-defined fully characterized gene, is often not straightforward. Splicing mutations can be difficult to gauge; apparently harmless synonymous substitutions can be pathogenic; missense mutation effects are not easy to predict (loss of function, gain of function, or no clear effect?).

Even for nonsense and frameshifting mutations, the effects of the mutation may be hard to predict. Largely depending on where a premature termination codon is introduced within the mRNA, the effect may be to trigger mRNA destruction, with failure to make a protein, or to produce a mutant protein that may or may not have a gain of function (see Box 7.1). And who would have expected that deleting a single nucleotide in the 2.5 Mb dystrophin gene could produce the severe Duchenne form of muscular dystrophy, while deleting a 1 Mb region containing that same nucleotide along with a million others (including many coding exons) would result in a much milder form of muscular dystrophy? Part of the explanation lay in the differential effects of frameshifting and in-frame deletions, and the observation prompted a novel RNA therapy for Duchenne muscular dystrophy, as described below.

For individuals affected by an autosomal recessive disorder there is an added complication: the need to assess the combined effect of two mutant alleles. In such

situations, the mutant alleles are typically loss-of-function mutations and the degree of overall residual function is the major determining factor.

For some disorders, such as many enzyme deficiencies, there is a good correlation between product levels and severity of the phenotype. In steroid 21-hydroxylase deficiency, for example, individuals with non-classical forms (later onset, mild) typically have 10–15 % of residual enzyme activity, whereas in classical forms (congenital, severe) there is from about 2 % residual enzyme activity, which usually manifests as a "simple-virilizing" phenotype, to 0 % enzyme activity in the most severe ("salt-wasting") form (the clinical phenotypes are given at the beginning of Clinical Box 6, on page 201). Phenotypes due to deficient X-linked hypoxanthine guanine phosphoribosyltransferase activity also show significant correlation with the amount of residual enzyme activity (Figure 7.22).

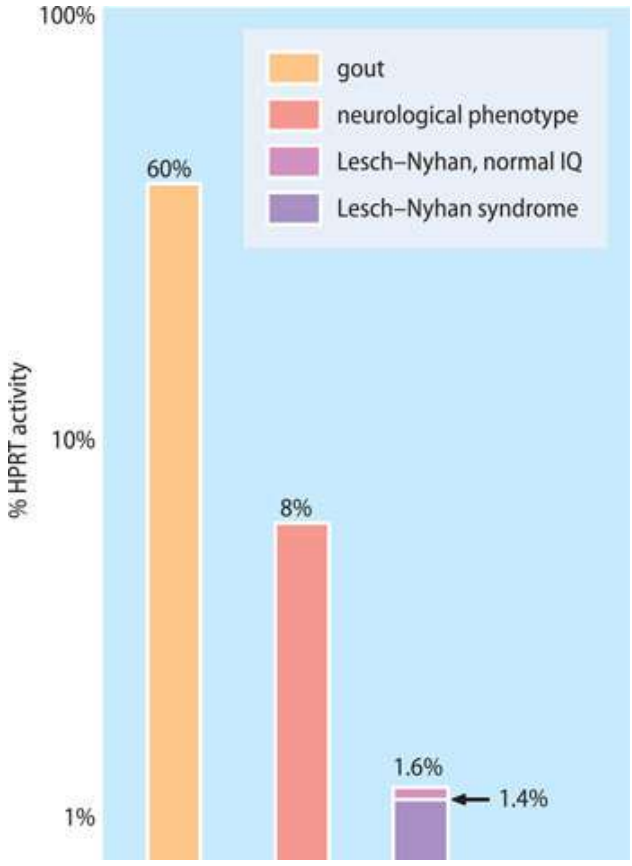


Figure 7.22 Different threshold levels for different phenotypes resulting from loss of activity of hypoxanthine guanine phosphoribosyltransferase (HPRT). Loss-of-function mutations in the X-linked *HPRT* gene can result in gout (which becomes manifest at less than 60 % of normal HPRT activity). If the HPRT activity falls to below 8 %, additional neurological features can begin to develop and are manifested as clumsiness and involuntary movements such as migrating

contractions (chorea) and twisting and writhing (athetosis). A decrease in HPRT activity to less than 1.4 % results in full Lesch–Nyhan syndrome (with choreathetosis and additional spasticity, self-mutilation, and mental retardation), but with an HPRT activity of about 1.4–1.6 %, individuals with Lesch–Nyhan syndrome can have normal intelligence.

Exceptional versus general reasons for poor genotype-phenotype correlations

For many monogenic disorders, genotype–phenotype correlations can be extremely complicated. Sometimes affected individuals who have identical mutant alleles (affected members of the same family; genotyped affected individuals within a population) show remarkable differences in phenotype.

We have already considered some exceptional factors that contribute to poor genotype–phenotype correlations in some Mendelian disorders: epigenetic factors (notably, parent-of-origin effects, as described in Section 6.3); dynamic mutations (due to an unstable expansion of short tandem repeats that can cause intergenerational differences in phenotype); and mosaicism (including differential X-chromosome inactivation that results in different effects for an X-linked mutation in women).

Disorders caused by mitochondrial mutations also show particularly poor genotype–phenotype correlation. Here, due to the exceptional problem of multiple copies of mtDNA per cell can be *homoplasmic* for a mutant allele (all mtDNA copies carry the mutation) or be *heteroplasmic* (with a mix of normal and mutant mtDNA molecules). Because egg cells typically contain more than 100 000 mtDNA molecules, every child of an affected heteroplasmic mother will inherit at least some mutant mtDNA molecules, but the proportion is difficult to predict and the ratio of mutant to normal mtDNA copies can change over time. As a result, mutations in mtDNA can have low penetrance and rather unpredictable effects.

Modifier genes and environmental factors: common explanations for poor genotype–phenotype correlations

In addition to the exceptional factors described above, two *general* factors can explain differences in the phenotype of a monogenic disorder in affected members of the same family (who can almost always be expected to have identical mutations in

the case of a monogenic disorder) and in affected individuals within a population who have been revealed to have identical genotypes.

One of these is genetic variation at other loci. Genes that interact with the disease locus to modify the disease phenotype are known as **modifier genes** (the interaction between a disease locus and a modifier gene locus is called *epista- sis*). Different alleles at a modifier locus can have different effects on a disease phenotype—they may sometimes have a protective effect (resulting in a milder disease phenotype) or an aggravating effect (inducing a more severe phenotype). The second general factor that influences a disease phenotype comes from the environment, as described below.

Modifier genes: the example of β-thalassemia

Until recently, modifier genes were not easy to identify directly in humans. Instead, heavy reliance was often placed on carrying out various types of analyses in animal disease models, which could then suggest candidate modifier genes for human diseases. Here we consider how modifier genes can affect the phenotype of a well-studied blood disorder, b-thalassemia.

Individuals with b-thalassemia have a genetic deficiency in b-globin, a component of hemoglobin. Although monogenic, this disorder is far from simple. It is usually autosomal recessive, but in occasional individuals the phenotype is dominantly inherited (one allele is normal; the other has an exceptional gain-of-function mutation). Although mutation in the b-globin gene, *HBB,* is the predominant factor in causing the disease, affected individuals with identical *HBB* alleles can show very significant differences in phenotype. Genetic variation at several modifier loci is also very important.

Adult hemoglobin is a tetramer with two a-globin chains and two b-globin chains; the synthesis of a- and b-globin chains is normally tightly regulated to ensure a 1:1 production ratio. However, when mutation in *HBB* results in a reduced production of b-globin chains, there will be a relative excess of a-globin chains. The excess a-globin monomers, present at high concentration, aggregate and precipitate, causing the death of early hemoglobin-producing cells in the bone marrow, and ineffective production of red blood cells. Those red blood cells that reach the peripheral blood also contain excess a-globin, which induces the formation of inclusion bodies and increased production of reactive oxygen species, leading to membrane damage and

hemolysis. Because the anemia that results from lower numbers of red blood cells is life-threatening, current therapy is based largely on blood transfusions.

Genetic variation at other globin loci can affect the clinical severity of b-thalassemia. Thus, a mutation causing a reduced output of a-globin chains reduces the globin chain imbalance and allows the production of more red blood cells. Normal individuals usually have two tandemly repeated a-globin genes (*HBA1* and *HBA2*) on each chromosome 16, but as a result of unequal crossover the number of copies of the a-globin gene can vary: 0 (–); 1 (–a); 2 (aa); 3 (aaa); or 4 (aaaa). Large numbers of a-globin genes can further add to the excess of a-globin chains that results from reduced b-globin production (**Figure 7.23**). As evidence of the modifier effect, individuals who are heterozygous for a null b-thalassemia (b^0) allele but have a total of six or more a-globin genes (aaa/aaa or aa/aaaa) can have a disease phenotype that resembles homozygous b-thalassemia).

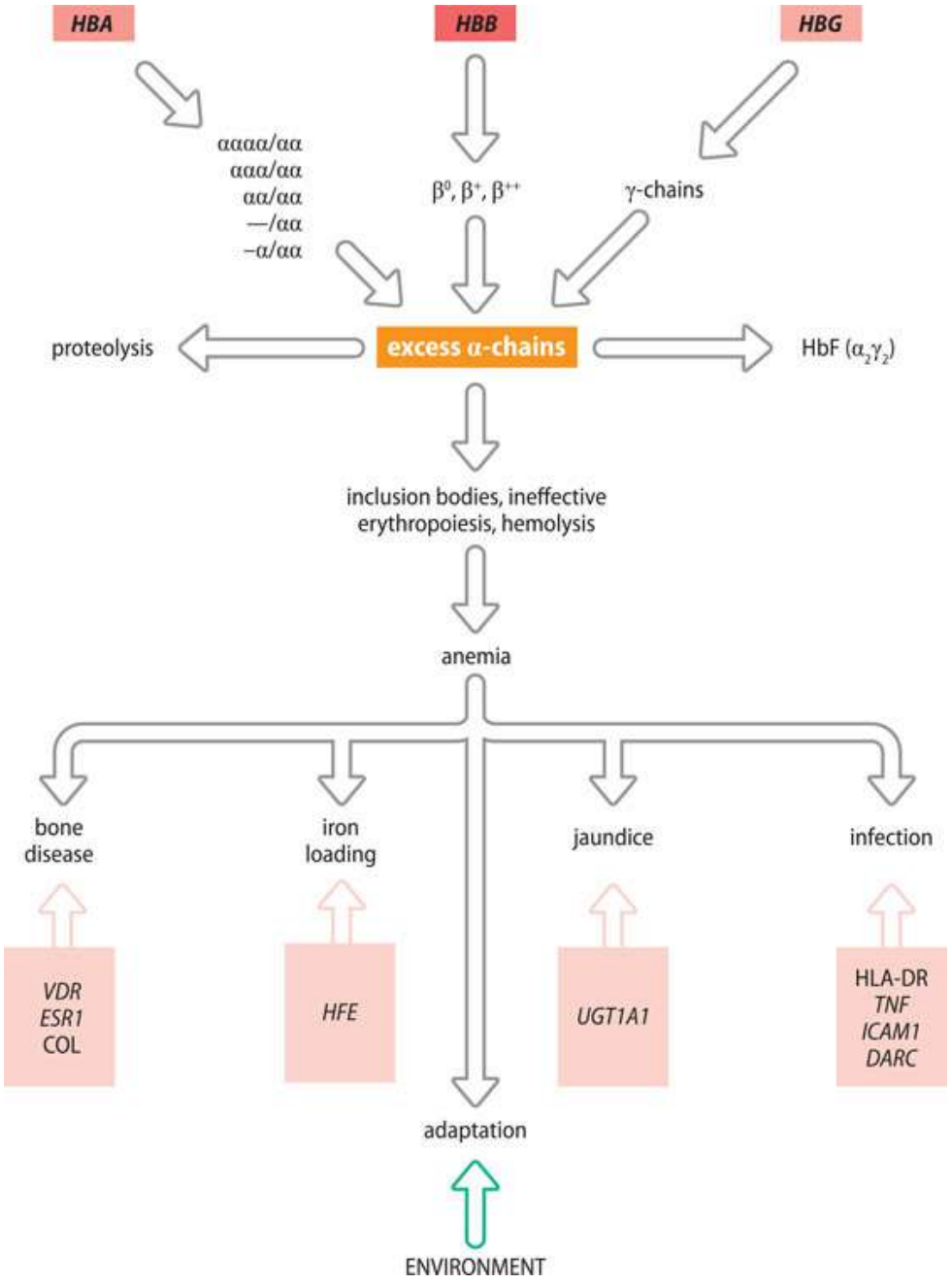


Figure 7.23 Multiple factors determine the β-thalassemia phenotype. The basic pathology of b-thalassemia results from excess a-globin chains damaging red blood cell precursors and red blood cells. Depending on the mutations at the disease locus *HBB,* there can be different levels of decrease in b-chain production (b^0, null; b^+, partial function; and so on) with a direct effect on the amount of excess a-globin chains. The phenotype can also vary as a result of variation in a-globin copy number, variation in the ability to produce HbF after birth (which uses up variable amounts of a-chains), and, possibly, through different rates of removal of a-chains by proteolysis. The many complications of the resulting anemia can also be modified by genetic variability, including

variation in the genes listed at the bottom. Genes or gene loci are: *HBB,* b-globin; *HBA*, a-globin loci; *HBG*, g-globin loci; *VDR,* vitamin D receptor; *ESR1,* estrogen receptor-1; COL, collagen loci; *HFE,* locus for hereditary hemochromatosis; *UGT1A1,* UDP glucuronyltransferase involved in bilirubin metabolism; HLA-DR, major histocompatibility complex loci; *TNF,* tumor necrosis factor; *ICAM1*, intercellular adhesion molecule-1; *DARC,* Duffy antigen receptor for chemokines. (Adapted from Weatherall D [2010] *Nature Med* 16:1112–1115. With permission from Macmillan Publishers Ltd.)

The β-thalassemia phenotype is also modified by genetic variants that control the production of hemoglobin F (HbF), which is composed of two a-globin chains and two g-globin chains. HbF is the dominant hemoglobin made during the fetal period (its high O_2-binding capacity makes it suited to working at the fetal stage), but there is a rapid decrease in HbF production at birth, and although it is still present at significant levels in infants it usually accounts for less than 1 % of hemoglobin in adults. However, HbF can account for 10–40 % of the hemoglobin in rare affected individuals with hereditary persistence of fetal hemoglobin, and there is significant variation between normal individuals in HbF levels. By forming more HbF, g-globin polypeptides compensate for reduced production of b-globin. The elevated HbF levels in infants is thought to be protective, explaining the delayed onset of symptoms in b-thalassemia, and comparatively high HbF levels at later stages may be partly protective.

Many of the complications of the disease are also modified by genetic variation at other loci (see Figure 7.23). There are also differences in the patterns of adaptation to anemia at different ages, and environmental factors, notably exposure to malaria, can also modify the phenotype.

Environmental factors influencing the phenotype of genetic disorders

In some disorders, expression of the disease phenotype depends very significantly on environmental factors that may act at different levels: at a distance (external radiation sources); by direct exposure of our cells to harmful or potentially harmful chemicals that we ingest (in food and drink) or inhale (such as tobacco smoke or atmospheric pollution); and by contact with microbes and toxins.

Especially important in triggering cancers, as described in Chapter 10, environmental factors are also very important in other complex diseases, whether at the earliest stages of development (factors in the uterine environment) or at later

stages (such as exposure to chemicals and microbes). We consider some aspects in Chapter 8.

Environmental factors are also known to be important in some single gene disorders. We illustrate the example of how dietary factors can influence disease with reference to phenylketonuria in **Clinical Box 9**. In Chapter 9, we also consider how differential sensitivity to drugs can influence other monogenic disorders, within the broader context of pharmacogenetics.

CLINICAL BOX 9 DISEASE PROFILE: PHENYLKETONURIA AS AN INBORN ERROR OF METABOLISM, A MULTIFACTORIAL CONDITION, AND AN EMBRYOFETOPATHY

The first genetic disorders that were investigated at the molecular level were *inborn errors of metabolism*. Affected individuals lacked a single enzyme that catalyzed one step in a metabolic pathway (usually consisting of a series of enzyme-catalyzed steps in which the product of one step becomes the substrate for the next step). Deficiency in one such enzyme would cause a metabolic block (**Figure 1A**). The resulting buildup of the substrate proximal to the block might drive an alternative pathway (red arrow). By analyzing blood and urine samples, pioneers in the field were able to obtain molecular clues as to the cause of a genetic disorder many decades before we knew about DNA structure and were able to study genes.

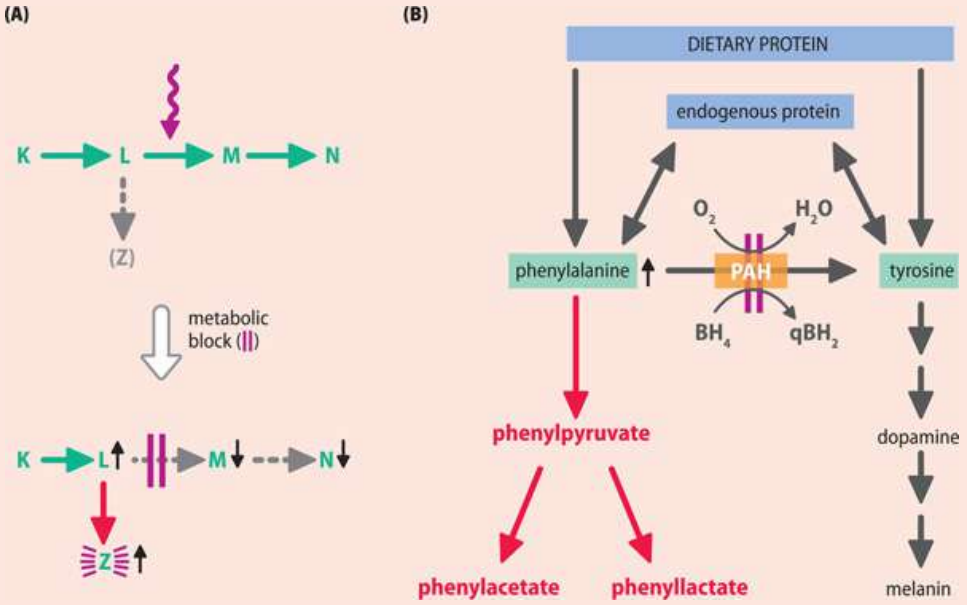


Figure 1 Metabolic blocks: principles and the example of phenylketonuria. (A) Principle of a metabolic block. Metabolites K, L, M, and N are linked by a series of enzyme-catalyzed reactions (green arrows) in which the product of an enzyme step serves as the substrate for the next enzyme. Here, as a result of genetic deficiency, there is a lack of the enzyme that converts L to M, leading to low concentrations of M, with a knock-on effect for the next step and a decreased concentration of N. The substrate L, proximal to the block, increases in concentration and that may lead to excessive production of metabolite Z, which becomes a biomarker of the disease. (B) Phenylalanine is converted to tyrosine by phenylalanine hydroxylase (PAH), which requires the cofactor tetrahydrobiopterin (BH4). When mutations cause homozygous deficiency in PAH, the conversion of phenylalanine to tyrosine is blocked (double magenta bar). As a result, high levels of phenylalanine build up (hyperphenylalaninemia), driving the production of new phenylalanine metabolites (shown by red arrows). Three phenylketones are produced (phenylpyruvate, phenylacetate, and phenyllactate) and are excreted. Deficiency of different genes involved in BH4 metabolism can also cause hyperphenylalaninemia.

Phenylketonuria was one of the earliest inborn errors of metabolism to be studied; it results from a deficiency of phenylalanine hydroxylase, a liver enzyme that converts phenylalanine to tyrosine (Figure 1B). Genetic deficiency in this enzyme results in elevated levels of phenylala-nine (hyperphenylalaninemia) that can be sub-clinical (120–600 μmol/liter) or in untreated individuals result in mild phenylketonuria (600-1200 μmol/liter) or classical phenylketonuria (>1200 μmol/liter).

Elevated phenylalanine concentrations drive the production of phenylketone derivatives (see Figure 1B), which are excreted. The clinical symptoms of phenylketonuria are largely due to the toxic effects of very high phenylalanine levels in the brain—untreated children show progressively impaired brain development, leading to severe intellectual disability and various other symptoms including behavioral problems.

The standard treatment is really a form of prevention. Infants identified as having very high levels of blood phenylalanine are placed on a low-phenylala-nine diet that is generally successful (although there may be problems with compliance in later years). The low-phenylalanine diet works because phenylketonuria is really a multifactorial disorder—two factors are absolutely required for the disease to manifest itself: a genetic factor (mutations at the *PAH* locus causing homozygous deficiency of phenylalanine hydroxylase) and an environmental factor (normal l-phenylalanine levels in dietary protein).

Phenylketonuria is classified as a monogenic disorder only because the vast majority of us are exposed to the environmental factor. Because affected sibs can show significant differences in clinical pheno-type, modifier genes are also likely to be involved. The phenotype could be influenced by genetic variation in different processes (such as protein degradation, phenylalanine transport and disposal, transport of phenylalanine across the blood–brain barrier, brain sensitivity to phenylalanine toxicity).

Very high levels of phenylalanine in phenylketonuria can be teratogenic and can result in *embryofetopathy.* A homozygous mother (who might nevertheless have a mild phenotype that could go unrecognized) can have heterozygous offspring who go on to develop mental retardation. During pregnancy the placenta naturally selects for higher concentrations of amino acids; as a result, phenylalanine levels may double in fetal blood, causing serious damage to brain and some other organ systems during development. Again, this can be prevented or ameliorated if the expectant mother is placed on a low-phenylalanine diet from the earliest stages of pregnancy.

SUMMARY

- A small fraction of genetic variation causes disease, either by altering the amount of gene products (via a change in gene copy number or gene regulation, or by introducing premature termination codons), or by changing the sequence of gene products.
- The genetic code is redundant (most amino acids can be specified by multiple different codons), and universal for nuclear genes; mitochondria use a slightly different genetic code.
- Synonymous single nucleotide substitutions (silent mutations) replace one codon by another without changing the amino acid. They occasionally cause disease by altering RNA splicing.
- Nonsynonymous substitutions replace an amino-acid-specifying codon by a codon specifying a different amino acid (missense mutation) or by a stop codon (nonsense mutation).
- A missense mutation is likely to be pathogenic if the replacement amino acid is physiochemically rather different from the original amino

acid.

- Splicing mutations often alter important splice junc tion sequences. Additional splicing mutations change other important splice regulatory sequences within exons and introns, or activate a cryptic splice site to make a novel splice site.
- Insertions and deletions produce a translational frame shift if the resulting number of coding sequence nucleotides is not exactly divisible by three. Such mutations usually introduce an in-frame premature stop codon. RNA splicing mutations can also cause a translational frameshift.
- An in-frame premature termination codon often sig nals degradation of the mRNA (nonsense-mediated decay), but if the premature stop codon is close to the normal stop codon a truncated protein is usually produced that may sometimes result in a more severe phenotype.
- In vertebrate DNA, the CG dinucleotide is a target for cytosine methylation and a hotspot for C ® T mutations; the resulting 5-methylcytosine is prone to deamination to give a thymine.
- DNA strands in a helix often pair out of register at runs of a short tandem repeat; replication in the mispaired region can cause pathogenic frameshifting insertions and deletions.
- Long arrays of CAG triplet repeats in coding DNA and various types of short tandem repeats in noncoding DNA can undergo unstable expansion. Such dynamic mutations show meiotic and mitotic instability and can cause disease by producing harmful proteins or RNAs.
- Nonallelic homologous recombination usually means a sequence exchange that occurs after pairing of nonallelic repeats with highly similar sequences.
- Reciprocal exchange between mispaired tan dem repeats on chromatids (unequal crossover) or between distantly spaced direct repeats on the same DNA molecule or on paired chromatid DNAs can result in deletions or duplications. In alternative nonreciprocal exchanges the sequence of one copy is replaced in part by the sequence of another copy (gene conversion).
- Exchange between inverted repeats on the same strand can produce pathogenic inversions.

- Breaks in the DNA of one chromosome can result in subchromosomal deletions, inversions, and also ring chromosomes (formed after a chromosome has lost a terminal segment on each arm and the two broken ends of the centromere-containing fragment join up).
- Translocations occur when two chromosomes undergo breakages and then exchange fragments. A balanced translocation means that there has been no obvious net loss of DNA.
- Aneuploidy involves the gain or loss of whole chro mosomes. The effects on the phenotype are due to a minority of genes that are especially dosage-sensitive.
- Disorders due to mtDNA mutations are maternally inherited and show variable ratios of mutant DNA to normal mtDNA (heteroplasmy). Clinical symptoms appear after heteroplasmy passes a threshold value that causes tissue damage due to defective oxidative phosphorylation.
- A mitochondrial genetic bottleneck occurs naturally in certain egg cell precursors, causing random severe reduction in the amount of mtDNA molecules passed to daughter cells. It can cause large variability in mutation load in egg cells produced by a heteroplasmic woman.
- Loss-of-function mutations can result in a null allele (complete absence of product, or complete functional inactivity) reduced expression, or an altered product with reduced functional activity.
- Haploinsufficiency means that a loss of function of one allele causes a phenotype in the presence of a working normal allele.
- Dominant-negative mutations result in a mutant gene product that somehow impairs the activity of the normal allele in a heterozygote.
- Gain-of-function mutations have a phenotypic effect in the presence of a normal allele that cannot be compensated for by producing more of the normal gene product.
- Loss-of-function and gain-of-function mutations in one gene can result in different phenotypes.
- Different components of a phenotype may be mani fested at different threshold levels of gene function.
- Prion proteins and related proteins, can misfold to give a structure prone to self-aggregation that can induce other normally folded

versions of the protein to misfold, thereby seeding protein aggregation to cause disease.

- Predicting the phenotype from the genotype is often difficult, even for a monogenic disorder. The effect of some types of mutation can be difficult to predict, and the phenotype is often influenced by genetic differences at other gene loci (modifier genes) and by environmental factors.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

Amino acid substitutions and silent mutations

Betts MJ & Russell R (2007) Amino-acid properties and consequences of substitutions. In *Bioinformatics for Geneticists*, 2nd ed (Barnes MR ed.), pp 311–341. Wiley-Blackwell.

Boyko AR (2008) Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet* 4:e100083; PMID 18516229.

Grantham R (1974) Amino acid difference formulae to help explain protein evolution. *Science* 185:862–864; PMID 4843792. [Gives a matrix that quantifies the effect of amino acid substitutions.]

Ng PC & Henikoff S (2006) Predicting the effects of amino acid substitutions on protein function. *Annu Rev Gen Hum Genet* 7:61–80; PMID 16824020.

Sauna ZE & Kimchi-Sarfaty C (2011). Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet* 12:683–691; PMID 21878961.

Nonsense-mediated decay (NMD)

Bhuvanagiri M (2010) NMD: RNA biology meets human genetic medicine. *Biochem J* 430:365–377; PMID 20795950.

Holbrook JA (2004) Nonsense-mediated decay approaches the clinic. *Nat Genet* 36:801–808; PMID 15284851.

Splicing and regulatory mutations

Cartegni L (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 3:285–298; PMID 11967553.

Jarinova O & Ekker M (2012) Regulatory variations in the era of next-gen sequencing: implications for clinical molecular diagnostics. *Hum Mutat* 33:1021–1030; PMID 22431194.

Sterne-Weiler T (2011) Loss of exon identity is a common mechanism of human inherited disease. *Genome Res* 21:1563–1571; PMID 21750108.

Wang GS & Cooper TA (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* 8:749–761; PMID 17726481.

Mutation rates, mutation load, and noncoding DNA mutations

Conrad DF (2011) Variation in genome-wide mutation rates within and between human families. *Nat Genet* 43:712–714; PMID 21666693.

French JD & Edwards SL (2020) The role of noncoding variants in heritable disease. *Trends Genet* 36:880–891; PMID 32741549.

Goriely A & Wilkie AOM (2012) Paternal age effect mutations and selfish spermatogonial selection: causes and consequences for human disease. *Am J Hum Genet* 90:175–200; PMID 22325359.

Keightley PD (2012) Rates and fitness consequences of new mutations in humans. *Genetics* 190:295–304; PMID 22345605.

MacArthur DG (2012) A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 335:823–828; PMID: 22344438.

Makrythanasis P & Antonorakis S (2013) Pathogenic variants in nonprotein-coding sequences. *Clin Genet* 84:422–428; PMID 24007299.

Xue Y (2012) Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am J Hum Genet* 91:1022–1032; PMID 23217326.

Pathogenic unstable expansion of short tandem repeats

Hannan AJ (2018) Tandem repeats mediating genetic plasticity in health and disease. *Nat Rev Genet* 19:286–298; PMID 29398703.

Khristich AN & Mirkin SM (2020) On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J Biol Chem* 295:4134–4170; PMID 32060097.

Rodriguez CM & Todd PK (2019) New pathologic mechanisms in nucleotide repeat expansion disorders. *Neurobiol Dis* 130:104515; PMID 31229686.

Gene conversion and nonallelic homologous recombination

Chen JM (2007) Gene conversion: mechanisms, evolution and human disease. *Nat Rev Genet* 8:762–775; PMID 17846636.

Liu P (2012) Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev* 22:1–10; PMID 22440479.

Chromosome abnormalities and nomenclature

McGowan-Jordan J, Hastings RJ & Moore S (eds.) An International System for Human Cytogenomic Nomenclature (2020) *Cytogenet Genome Res* 160: Issues 7–8.

Nagaoka SI (2012) Human aneuploidy: mechanisms and new insights into an age-old problem. *Nat Rev Genet* 13:493–504; PMID 22705668.

Roukos V (2013) The cellular etiology of chromosome translocations. *Curr Opin Cell Biol* 25:357–364; PMID 23498663.

Weckselblatt B & Rudd MK (2015) Human structural variation: mechanisms of chromosome rearrangements. *Trends Genet* 31:587–599; PMID 26209074.

Molecular pathology of mitochondrial disorders

Craven L (2017) Recent advances in mitochondrial disease. *Annu Rev Genom Hum Genet* 18:257–275; PMID 28415858.

Gasparre G & Porcelli AM eds. (2020) *The Human Mitochondrial Genome – From Basic Biology to Disease*. Academic Press.

Gusic M & Porkisch H (2021) Genetic basis of mitochondrial diseases. *FEBS Lett* 595:1132–1158; PMID 33655490.

MITOMAP database at https://www.mitomap.org [a human mitochondrial genome database, with pathogenic mtDNA mutations and associated clinical characteristics etc].

Stewart JB & Chinnery PF (2021) Extreme heterogeneity of human mitochondrial DNA from organelles to populations. *Nat Rev Genet* 22:106–118; PMID 32989265.

Gain-of-function mutations, haploinsufficiency and molecular basis of genetic dominance

Lester HA & Karschin A (2000) Gain-of-function mutants: ion channels and G protein-coupled receptors. *Annu Rev Neurosci* 23:89–125; PMID 10845060.

Veitia RA & Birchler JA (2010) Dominance and gene dosage balance in health and disease: why levels matter! *J Pathol* 220:174–185; PMID 19827001.

Wilkie AOM (1994) The molecular basis of genetic dominance. *J Med Genet* 31:89–98; PMID 8182727.

Protein misfolding and protein aggregation in disease

Chiti F & Dobson CM (2017) Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu Rev Biochem* 86:27–68; PMID 28498720.

Cox D (2020) Protein aggregation in cell biology: an aggregomics perspective of health and disease. *Semin Cell Dev Biol* 99:115–130; PMID 29753879.

Dobson CM (2020) The amyloid phenomenon and its significance in biology and medicine. *Cold Spring Harb Perspec Biol* 12:a033878; PMID 30936117.

O'Carroll A (2020) Prions and prion-like assemblies in neurodegeneration and immunity: The emergence of universal mechanisms across health and disease. *Semin Cell Dev Biol* 99:115–130; PMID 31818518.

Modifier genes

Drumm ML (2012) Genetic variation and clinical heterogeneity in cystic fibrosis. *Annu Rev Pathol* 7:267–282; PMID 22017581.

Riordan JD & Nadeau JH (2017) From peas to disease: modifier genes, network resilience and the genetics of health. *Am J Hum Genet* 101:177–191; PMID 28777930.

Sankaran VG (2010) Modifier genes in Mendelian disorders: the example of hemoglobin disorders. *Ann NY Acad Sci* 1214:47–56; PMID 21039591.

8

Identifying disease genes and genetic susceptibility to complex disease

DOI: 10.1201/9781003044406-8

CONTENTS

Until recently, the molecular identification of rare genes for monogenic disorders was laborious, often consuming many years of painstaking effort to identify even just one disease gene. Now, in the era of massively parallel DNA sequencing (next-generation sequencing), it has almost become routine. We cover the principles in Section 8.1. Some diffi culties remain, however, because for some single gene disorders the disease phenotypes do not have a very well-defined, distinctive pathology. And a good deal of follow-up work will be needed to dissect out all the factors in monogenic diseases, which are sometimes rather complex, as described in Section 7.9.

The next big challenge has been to identify genes underlying complex (multifactorial) diseases, in which there is no obviously predominant disease locus (at least, not to the extent found in monogenic disorders). Instead, expression of the disease phenotype may be notably dependent on a few genes (oligogenic disorders) or many genes (polygenic disorders), with variable (and sometimes very strong) contributions from environmental factors. We cover the background to complex disease and polygenic theory in Section 8.2.

The genetic contribution to complex diseases differs according to the disease and between populations, and its overall impact can vary within a single population, depending on changeable environmental conditions. Investigations into the genetic susceptibility to multifactorial disease began decades ago but had limited success until the early 2000s. More recently, however, many DNA variants have been identified to confer susceptibility to complex diseases (genetic risk factors) and some have been shown to lower disease susceptibility (protective factors). We outline the different general approaches used to uncover the genetic susceptibility to complex diseases in Section 8.2.

In Section 8.3 we consider selected aspects of the genetic architecture of complex disease to illustrate the progress that has been made. We then follow up by considering gene-environment interactions and the contribution made by epigenetic factors. We consider investigations of common cancers separately in Chapter 10.

8.1 IDENTIFYING GENES IN MONOGENIC DISORDERS

A historical overview of identifying genes in monogenic disorders

Identifying genes underlying monogenic disorders began with a very few exceptional cases in the 1970s and early 1980s. The underlying genes were able to be identified through a known protein product (*functional cloning*), or as a result of huge enrichment of corresponding mRNAs in certain cell types. In the former case, for example, hemophilia A was known to be due to a deficiency of blood clotting factor VIII. Enough factor VIII protein was purified from pig blood to obtain a partial amino acid sequence. After an optimal short sequence of amino acids was identified, a panel of different but related oligonucleotides was synthesized to cover all codon possibilities for the selected sequence of amino acids. The resulting oligonucleotides were used as probes to screen DNA libraries, identifying first homologous cDNA clones and eventually human gene clones.

An alternative approach was candidate gene testing. Candidate genes could be selected on the basis of knowing the biology of the condition or often from observed similarities between the disease phenotype and a highly related phenotype in humans or animals. After the *FBN1* fibrillin gene was shown to be a locus for Marfan syndrome, for example, the related *FBN2* fibrillin gene was quickly and successfully investigated as a candidate gene for a very similar disorder, congenital contractual arachnodactyly. And after the mouse *Sox10* gene was identified as the locus for the *Dominant megacolon* (*Dom*) phenotype, a mouse model of Hirschsprung disease, human *SOX10* was quickly shown to be mutated in Waardenburg–Hirschsprung disease.

Positional cloning strategies were needed to identify genes underlying diseases where little was known about the kind of gene product they might make. They relied on first getting a subchromosomal position for the disease gene. (For X-linked conditions, at least the location had already been narrowed down to a single chromosome.) As listed below, two

types of approach were used to identify subchromosomal locations for an underlying gene in monogenic disorders.

- *Linkage analysis*. This is a general method applicable to the great majority of monogenic disorders. Blood samples would be collected from multiple family members for each of many families with the disorder. DNA from the blood samples would then be used to assay genotypes for each of a collection of hundreds of DNA markers that had previously been mapped to specific subchromosomal areas from across the genome. If a particular DNA marker co-segregated with disease, the disease gene could be inferred to map to the same subchromosomal location as the marker.
- *Chromosome break mapping*. Linkage analysis is not suitable for some disorders, notably dominant disorders where affected individuals fail to reproduce. Scanning patient blood samples for chromosome breaks can be profitable because of the high chance that *de novo* chromosome breaks (from translocations, deletions, or inversions) could be disease-associated (an important gene might have been inactivated as a result of the chromosomal damage; if so, it must lie within the deleted area or close to a chromosome breakpoint).

Positional cloning versus positional candidate approaches

Early efforts at getting a subchromosomal position for a monogenic disorder ran into the problem that very little would be known about the DNA sequences and genes in that region; DNA sequences from the same subchromosomal area needed to be cloned, sequenced, and tested. Such positional cloning efforts could be hugely laborious. Interested readers can get an idea of what was involved in positional cloning of the cystic fibrosis gene from PMID 23378595.

The task became easier as the genome began to be deciphered, with data from multiple laboratories being used to map large numbers of protein-coding genes to subchromosomal regions that had previously been poorly studied. As a result, much of the slog of cloning DNA was avoided, and the projects became *positional candidate approaches*. Genes in the subchromosomal region of interest could now be identified by consulting gene, genome, and literature databases, further studied as required, and then prioritized for mutation screening, according to their known characteristics. Ultimately, detailed gene maps were obtained for each chromosome; they were enormously helpful for positional candidate approaches with which to identify disease genes.

The final step: mutation screening

The final step is to scan DNA samples from affected individuals for disease-associated mutations in candidate genes (by comparison with unaffected controls). For a highly penetrant dominant disorder, pathogenic mutations should normally be found in affected individuals only; for a recessive disorder, the pathogenic mutations will occasionally be found in normal individuals (who might then be suspected to be heterozygous carriers). The list of candidate genes mapping in the relevant subchromosomal region might often be daunting, and computer programs were developed to prioritize the most likely candidate disease genes.

Mutation screening typically involves amplifying individual exons and the immediately surrounding intronic sequence from individual candidate genes and sequencing them to identify mutations associated with disease in panels of DNA samples from affected individuals and controls. As we will see in the final part of Section 8.1, however, genome-wide gene or exon sequencing can often dispense with the time-consuming need to first identify candidate disease genes.

Protein-coding genes have been the overwhelming choice for candidate disease genes. Loss-of-function mutations are expected in recessive conditions, and in dominant disorders resulting from haploinsufficiency. They are relatively easy to identify in protein-coding genes because they often occur in coding DNA sequences (where it is easy to spot mutations causing premature termination codons, changes to the reading frame, or amino acid substitutions) or close to exon–intron boundaries (causing splicing abnormalities). Gain-of-function mutations often involve specific missense mutations (which would not be expected in controls). A tiny number of RNA genes have, however, been identified as loci for monogenic disorders (see Table 7.5 on page 191).

Linkage analysis to map genes for monogenic disorders to defined subchromosomal regions

Genetic markers (polymorphic loci) from across the genome can be used to track the inheritance of a gene by using **linkage analysis**. Different types of linkage analysis can be carried out, but success usually depends on having suitably informative families with multiple affected individuals. Because human family sizes are generally very small, multiple different families need to be investigated. Hundreds of genetic markers are needed, from defined locations distributed across the genome. That became possible with the development of human genetic maps.

Human genetic maps

Human genetic mapping has been a recent endeavor, unlike genetic mapping in model organisms where gene mutations causing readily identifiable phenotypes can easily be mapped. In *Drosophila,* for example, crosses can be set up to breed mutant white-eyed flies

with flies that have abnormal curly wings; the progeny are then examined to see if the two mutant phenotypes segregate together or not. In humans, however, that kind of approach could never be applied—a different strategy was needed.

Instead of having a genetic map based on gene mutations, the solution to making a human genetic map was to identify *general* DNA variants (which rarely map within coding sequences, and usually have no known effects on the pheno-type). Different types of variants were identified and mapped to specific genome locations, beginning with restriction fragment length polymorphisms (RFLPs) that created or destroyed a restriction site (Figure 4.4), followed by microsatellite polymorphisms that varied in the copy number of short tandem repeats (Figure 4.5).

The first comprehensive map of human genetic markers (polymorphisms) did not appear until 1994. Based on microsatellite and restriction site polymorphisms, it had a marker spacing of just over one marker per megabase of DNA. Microsatellite markers have the advantage that they are highly polymorphic (with multiple alleles, each having a different number of copies of the repeat), whereas restriction site polymorphisms often have just two alleles.

The most recent maps are based on single nucleotide polymorphisms (SNPs). They also have limited polymorphism, with often just two alleles. But they have two very strong advantages: they are extremely abundant in the human genome, and they are amenable to automated typing.

Data on individual SNPs can be accessed at the dbSNP database (https://www.ncbi.nlm.nih.gov/snp/). Identifying reference numbers are composed of a seven-digit to nine-digit number prefixed by rs (= reference SNP), such as rs1800588. The database can also be queried with a gene symbol to find SNPs in a specific gene; the resulting data can be filtered progressively to get human SNPs, and then SNPs that are of clinical interest or that have been recorded in the corresponding locus-specific database.

Principle of genetic linkage

One fundamental principle underlies genetic linkage: alleles at very closely neighboring loci on a DNA molecule are co-inherited because the chance that they are separated by recombination is very low. By extension, alleles at distantly spaced loci on the same DNA molecule are much more likely to be separated by recombination at meiosis. (During human meiosis, chromosomes are often split by recombination into between two and seven segments—see **Figure 8.1** for an example.)

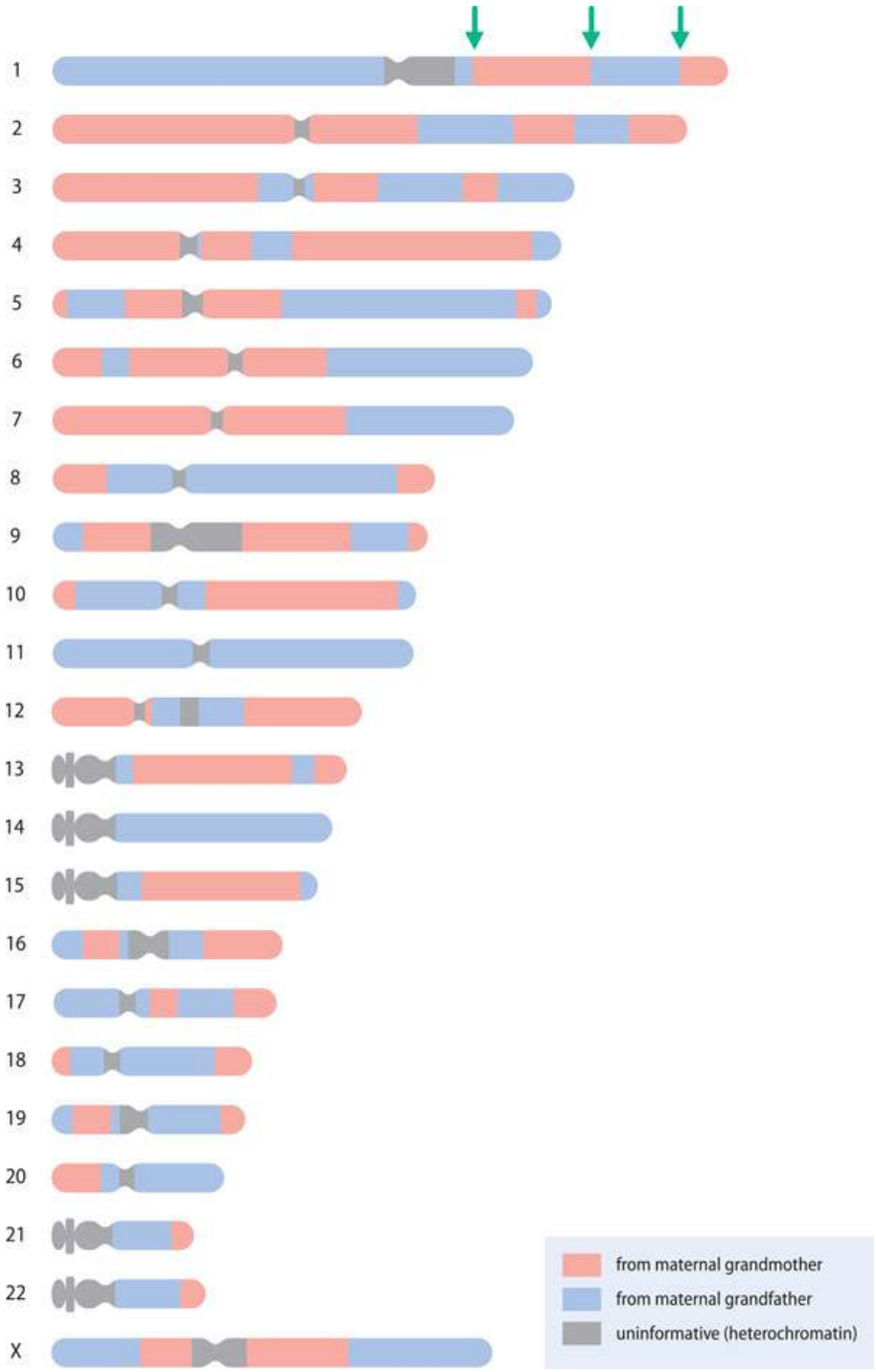


Figure 8.1 Mapping human meiotic crossovers by genome-wide SNP genotyping in families. This example shows the deduced pattern of meiotic recombination in the chromosomes passed on from a mother to her daughter, after whole genome SNP typing of the family members, including grandparents. The egg transmitted by the mother to her daughter was formed by meiotic cell division in which chromosomes from the maternal grandmother (pink) recombined with those from the maternal grandfather (blue) to give the alternating blue and pink patterns shown (as an illustration, the green arrows show the location of three

crossovers on chromosome 1). More than 50 crossovers can be detected, but none are apparent, unexpectedly, on chromosomes 11 and 14 (undetected crossovers might have occurred here in heterochromatic regions where no SNP markers were available; most of the short arm of chromosome 14 is composed of heterochromatin). (Courtesy of Rosemary J Redfield.) (CC BY-SA 2.5 CA).

A **haplotype** is a series of alleles at two or more neighboring loci on a *single* chromosomal DNA molecule. In human genetics, the term was first widely used within the context of the HLA system (readers who are unclear about haplotypes might wish to have a look at Figure 2 in Box 4.3 on page 106 to see how haplo-types are derived after obtaining genotypes from multiple family members).

We illustrate the principle of a *disease haplotype* in **Figure 8.2** within the context of a gene for an autosomal dominant disorder. The marker loci flanking the disease locus in Figure 8.2 are imagined to be very close to the disease locus. Recombination within this region would be extremely rare—the marker loci would be said to be *tightly linked* to the disease locus.

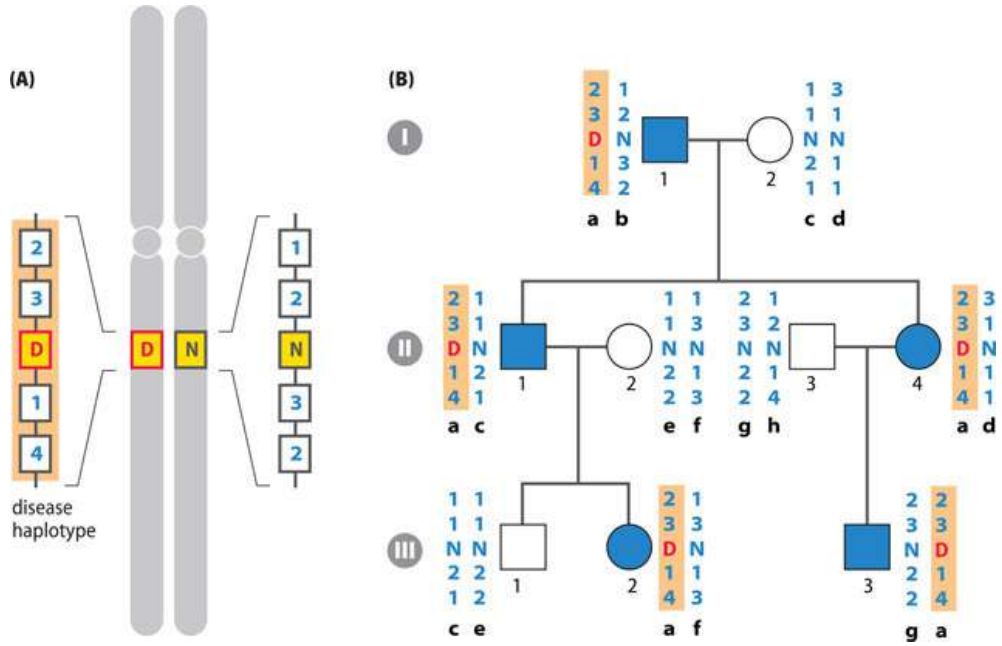


Figure 8.2 Inheritance of a disease haplotype in an autosomal dominant disorder. (A) The disease locus, highlighted by the yellow box, can be imagined to have two alleles: D (disease) and N (normal). Here we also consider alleles at two proximal marker loci and two distal marker loci that are physically located very close to the disease locus, for example within 0.5 Mb. The disease haplotype in the affected individual is defined by the sequence of *alleles* at consecutive neighboring marker loci that could be represented here as 2–3–1–4 (when read in the proximal to distal direction). (B) The haplotypes in (A) belong to the affected grandfather (I-1) in this pedigree. The highlighted disease haplotype (**a**) is transmitted without change to affected individuals in generations II and III.

In Figure 8.2B the disease haplotype is transmitted unchanged through four meioses: from the grandfather (I-1) to his affected son and daughter, and then to the two affected grandchildren (III-2 and III-3). For marker loci that are increasingly distant from the disease locus, the incidence of recombination between disease locus and marker locus becomes progressively greater.

As the distance separating a marker locus and the disease locus on the same chromosome increases, a point will be reached where the chance of recombination between the two loci would equal the chance of no recombination between them. The marker would then be said to be *unlinked* to the disease locus (it would be no different from a marker on a different chromosome, for which an allele would have a 50 % chance of segregating with disease, just by chance).

Human meiotic recombination frequencies

For any two loci, the chance of recombination is a measure of the distance between them. Loci separated by recombination in 1 % of meioses are said to be 1 *centimorgan* (*cM*) apart. Genetic distances are related to physical distances, but not in a uniform way: there is a rough correspondence between a genetic distance of 1 cM in humans and a physical distance of close to 1 Mb of DNA, but there are considerable regional variations across chromosomes.

Recombination is much more common at subtelomeric regions than in the middle of chromosome arms, for example, and is much less frequent in heterochromatic regions. At higher resolution, the majority (60 % or more) of crossovers occur at a number of short hotspots, about 1–2 kb long, across the genome.

Recombination frequencies in human meiosis also show significant sex differences. Using dense genome-wide SNP mapping in nuclear families, two large studies (by Cheung et al. [2007] and Coop et al. [2008]—see under Further Reading) looked at 728 and 557 meioses respectively; the overall mean scores averaged across the two studies was 25.2 crossovers in male meiosis and 39.1 crossovers in female meiosis. But there is also variation between individuals, and even between individual meioses within a single individual (as shown in **Figure 8.3**). Overall, therefore, there is no one correct human genetic map length: for any one subchromosomal region, the correspondence between the physical (DNA sequence) length and genetic map length will vary from one meiosis to another.

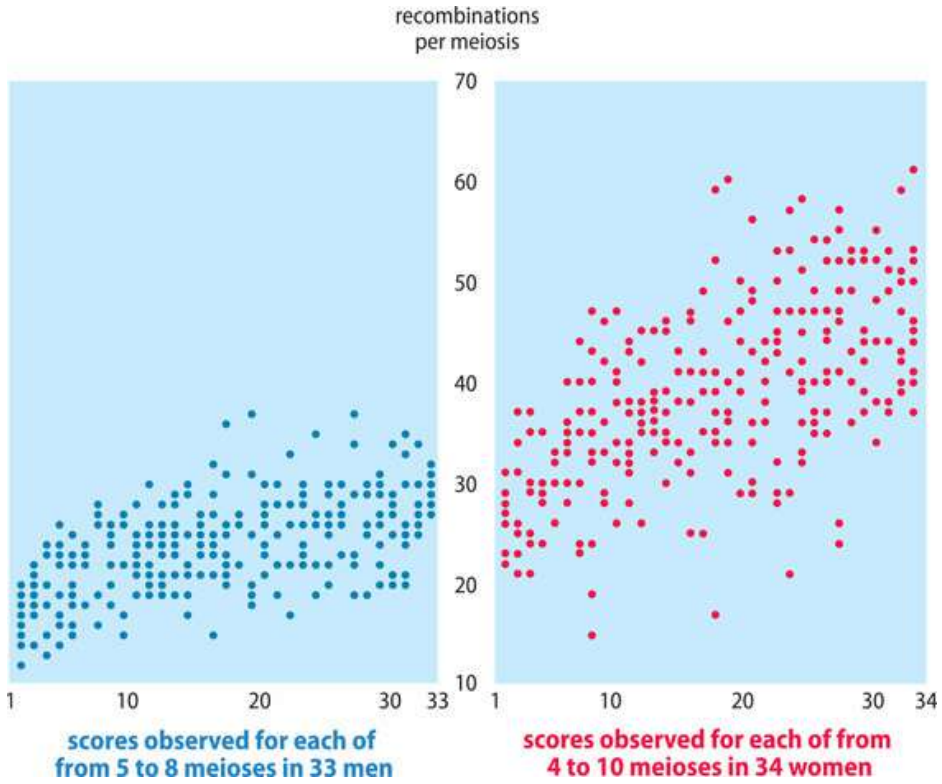


Figure 8.3 Individual differences in the numbers of recombinations per meiosis. Each dot represents the number of recombinations identified in an individual meiosis; each vertical line of dots represents the scores determined in each of multiple meioses in a single male or female, as shown. The number of recombinants per meiosis was determined by genotyping individuals in families with 6324 SNP markers. (From Cheung VG et al. [2007] *Am J Hum Genet* 80:526–530; PMID 17273974. With permission from Elsevier.)

Standard genome-wide linkage analyses

To map disease genes to specific chromosomal regions, genome-wide linkage analyses can be used. Usually, several hundred genetic markers from defined loci across the whole genome are genotyped in family members with the disease. The results may show some marker loci that are tightly linked to the disease, thereby indicating a subchromosomal location for the disease gene.

Using, say, 400 markers for genome-wide linkage analyses would give a marker density of one every 7–8 Mb or so. Given that our chromosomes are about 50–250 Mb in length and are split by meiotic recombination into usually only two to seven segments (Figure 8.1), there is a high chance that one or more of a 400 genome-wide marker set will be sufficiently close to the disease locus for a marker allele to co-segregate with a disease allele.

The segregation of alleles from each marker locus is followed through a suitably large number of informative meioses (see **Figure 8.4** for examples of informative and

uninformative meioses). In practice that means having access to samples from multiple affected and unaffected members usually drawn from several families.

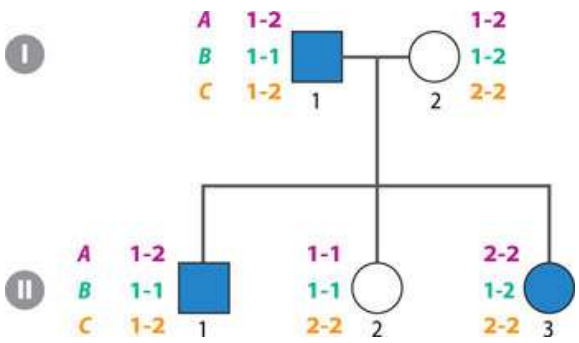


Figure 8.4 Informative and uninformative meioses. Let us assume full penetrance and autosomal dominant inheritance of the phenotype in the pedigree shown here. The disease allele has been transmitted from the father to the son and one daughter (II-3). Genotypes for three unlinked marker loci, *A, B,* and *C*, are shown by the respective colored figures. Consider marker *A*. For the affected son it is impossible to tell which parent contributed allele 1 and which contributed allele 2, but it is possible to infer that each parent contributed an allele 1 to II-2 and an allele 2 to II-3. Marker *B* is completely uninformative here because I-1 is homozygous and it is impossible to tell which of the two paternal alleles 1 was transmitted to each child. Marker *C* is informative in each case: the father transmitted allele 1 to his affected son, but transmitted allele 2 to both daughters.

In an idealized situation, **recombinants** and non-recombinants can be clearly identified. In the autosomal dominant pedigree in **Figure 8.5**, the affected individual in generation II is heterozygous for a disease allele and he is also heterozygous for a marker, having a 1,2 genotype. In the highly unusual circumstances shown in this figure it is possible to identify recombinants and non-recombinants unambiguously. In practice, recombinants often cannot be identified unambiguously (linkage studies often use families that do not have such an ideal structure, and key meioses may often be uninformative).

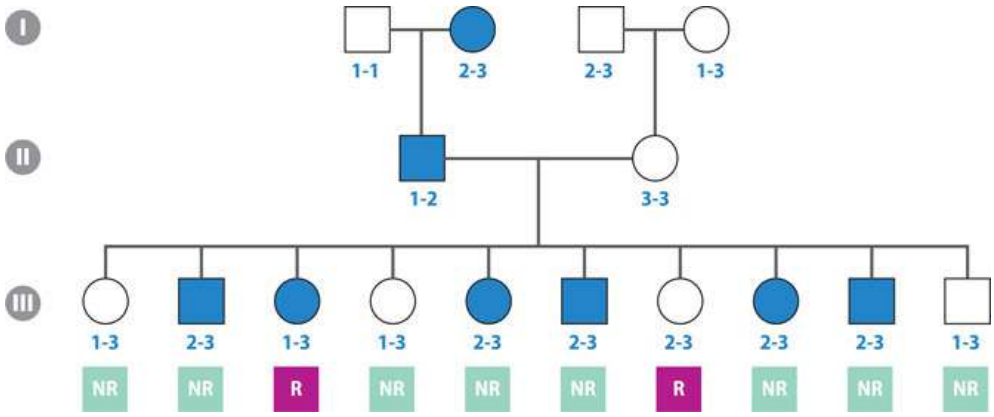


Figure 8.5 Unambiguous identification of recombinants and non-recombinants in an idealized pedigree. Members of this autosomal dominant pedigree have been typed for a marker that has three alleles (1,2, and 3).

The available data suggest that allele 2 of the marker is segregating with the disease. In that case, 8 out of the 10 children in generation III are non-recombinants (NR): either they have inherited both the disease and allele 2 from their affected father, or they have inherited paternal allele 1 and are unaffected. The other two are recombinants (R): either paternal allele 1 has segregated with disease, or paternal allele 2 is associated with a normal allele at the disease locus.

To get round the difficulties of identifying recombinants in human genetic mapping, sophisticated computer programs are needed. They do not attempt to identify individual recombinants. Instead, their job is to survey the linkage data and then calculate alternative probabilities for linkage and nonlinkage. They then express the ratio of these probabilities as a logarithmic value called a **lod score**, as described in **Box 8.1**. Programs such as these are dependent on previous information on the mode of inheritance, disease gene frequency, and the penetrance of the genotypes at the disease locus (that is, the frequency with which the genotypes manifest themselves in the phenotype). For monogenic disorders, the mode of inheritance and disease gene frequency often do not present much difficulty; penetrance can be a more difficult problem.

BOX 8.1 LOD SCORES AND STATISTICAL EVIDENCE FOR LINKAGE

Computer-based linkage analysis programs calculate two alternative probabilities: (i) the likelihood of the marker data, given that there is linkage between the marker and the disease locus at specified recombination fractions; and (ii) the likelihood of the marker data, assuming that the marker is unlinked to the disease locus. The *likelihood ratio* is the ratio of likelihood (i) and likelihood (ii), and provides evidence for or against linkage.

The convention has been to use the logarithm of the likelihood ratio, called the **lod score** (logarithm of the odds). Individual lod scores are calculated for a defined recombination fraction (θ), and so for each marker the computer programs provide a table of lod scores for different recombination fractions (θ = 0, 0.10, 0.20, 0.30, and so on). The reported recombination fraction is chosen to be the one where the lod score (Z) is at a maximum (Z_{max}).

A lod score of +3 is normally taken to be the *thresh- old* of statistical significance for linkage between two loci. It means that the likelihood of the data given that the two loci are linked is 1000 times greater than if it is unlinked ($\log_{10}1000 = 3$). Linkage between two loci (say, a disease locus and a marker locus) can theoretically be achieved with 10 informative meioses. At each informative meiosis there are two choices: the two loci are linked (a specific marker allele segregates with disease), or they are not linked (the marker allele does not segregate with disease). If the same marker allele segregates with disease in each of 10 informative meioses in a large pedigree, the odds of that happening by chance are 2^{10} to 1 against, or just over 1000:1 against. In practice, because of poor family structures and some uninformative meioses, 20 or more meioses are often needed for

linkage to be successful; DNA samples are usually needed from affected and unaffected members of multiple families.

The ratio of 1000:1 might seem overwhelming evidence in favor of linkage, but it is required to offset the inherent improbability of linkage. With 22 autosomes, two randomly chosen loci are unlikely to be on the same chromosome. Even if the two loci are on the same chromosome, however, they may be well separated, and so unlinked. Factoring in both of these observations, the prior odds are about 50:1 *against* linkage, or 1:50 in favor of linkage. That means we need pretty strong evidence from linkage analysis data to counteract the low starting probability. A likelihood ratio of 1000:1 in favor of linkage multiplied by a prior odds of 1:50 in favor of linkage gives a final odds of only 20:1 in favor of linkage. That is, a single lod score of 3 is not proof of linkage; there is a 1 in 20 chance that the loci are not linked.

Higher lod scores provide greater support for linkage. A lod score of 5 is 100 times more convincing than a lod score of 3. In practice, therefore, genome-wide claims for linkage based on a single lod score less than 5 should be treated as provisional evidence for linkage. However, significant lod scores may often be obtained for several markers clustered in one subchromosomal region; if so, the combined data provide strong evidence of linkage. See **Table 1** for the example of a dominantly inherited skin disorder, Hailey–Hailey disease (OMIM 169600), in which four neighboring markers in the 3q21-q24 interval show significant evidence of linkage.

TABLE 1 PAIRWISE LOD SCORES FOR HAILEY-HAILEY DISEASE AND MARKERS AT 3Q21-Q24

	Lod score (Z) at				Maximum likelihood estimates	
Marker	**0.00**	**0.10**	**0.20**	**0.30**	**Z_{max}**	**AT θ =**
D3S1589	−0.99	2.29	1.90	1.14	2.29	0.09
D3S1587	4.54	3.80	2.83	1.73	4.54	0.00
D3S1292	2.62	4.98	3.84	2.41	5.32	0.04
D3S1273	3.36	5.52	4.12	2.54	6.10	0.03
D3S1290	−2.81	3.83	3.05	1.94	3.90	0.07
D3S1764	−8.62	2.21	2.06	1.38	2.26	0.13

The descending order of markers is from proximal to distal. Analyses were carried out in six disease families. The underlying disease gene was subsequently found by positional cloning to be *ATP2C1* and to map just proximal to *D3S1587*. (Data from Richard G et al. [1995] *J Invest Dermatol* 105:357-360; PMID 7665912).

The threshold for excluding linkage is a lod score of −2. *Exclusion mapping* can be helpful in excluding a candidate gene of interest, and in genome-wide studies the exclusion of a substantial fraction of the genome can direct extensive analysis of the remaining regions.

Linkage can theoretically be achieved with 10 informative meioses, but in practice linkage analysis is rarely successful when there are fewer than 20 or so meioses (see Box 8.1). A major confounding problem in linkage analysis is locus heterogeneity, when the same disease in different families under study may be caused by different genes—it is important to try to study families with extremely similar disease phenotypes.

After obtaining evidence of linkage, crossover points are deduced to identify a minimal subchromosomal region for a disease gene (**Figure 8.6**).

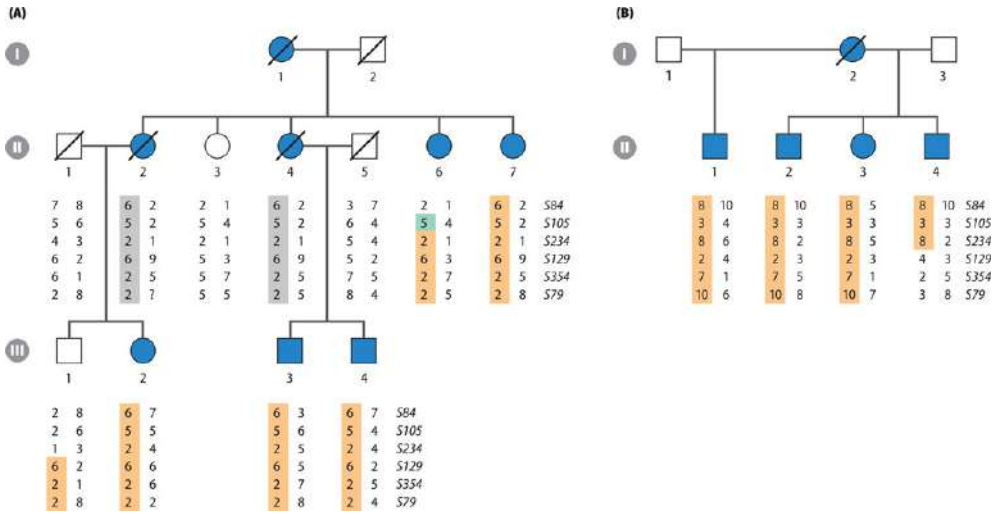


Figure 8.6 Defining the minimal candidate region by inspection of haplotypes. The two pedigrees show a dominantly inherited skin disorder, Darier–White disease, that had previously been mapped to 12q. The 12q marker haplotype segregating with disease is highlighted by orange shading. Gray boxes mark inferred haplotypes in deceased family members. In pedigree (A) the recombination in II-6 maps the disease gene distal to marker *D12S84* (abbreviated to *S84* in the figure); the *D12S105* marker is uninformative because I-1 was evidently homozygous for allele 5—compare the genotypes of II-3 and II-7. The recombination shown in III-1 suggests that the disease gene maps proximal to *D12S129,* but this requires confirmation (the interpretation depends on the genotypes of II-1 and II-2 being inferred correctly, and on III-1 not being a non-penetrant gene carrier). The recombination in II-4 in pedigree B provides the confirmation. The combined data locate the Darier gene to the interval between *D12S84* and *D12S129.* (Adapted from Carter SA et al. [1994] *Genomics* 24:378–382; PMID 7698764. With permission from Elsevier.)

Autozygosity mapping in extended inbred families

The term **autozygosity** means homozygosity for markers that are *identical by descent*—the two alleles are copies of one specific allele transmitted to both parents from a recent common ancestor. In some societies, such as in the Middle East and parts of Asia, cousin marriages are quite common, and in extended inbred families there may be several individuals who are autozygous for an allele because of parental consanguinity. As

illustrated in Figure 1 of Box 5.2 on page 115, second cousins share respectively 1/32 of their genes, and so their children would be autozygous at 1/64 of all loci.

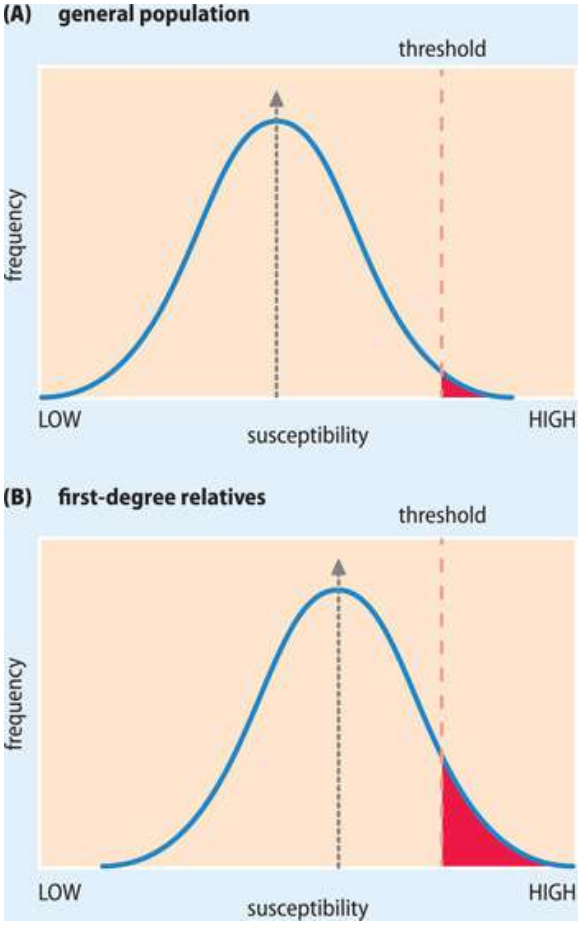


Figure 1 Distribution of susceptibility to a complex disease in the general population and in first-degree relatives of an affected person.

Homozygosity for a particular marker allele could be due to autozygosity. Alternatively, it might result from the inheritance of a second, independent copy of that allele that has been brought into the family at some stage; alleles such as this would be said to be *identical by state*. Homozygosity for a haplotype of marker alleles, however, is more likely to indicate autozygosity; if there are additional sibs who are homozygous for the same marker haplotypes, quite small consanguineous families can generate significant lod scores. Autozygosity mapping can therefore be a very efficient way of mapping a recessive monogenic disorder; see Goodship et al. (2000) under Further Reading for a successful application.

Chromosome abnormalities and other large-scale mutations as routes to identifying disease genes

Some affected individuals show a specific chromosome abnormality or other very large-scale mutation that can be detected quite readily (see below). Abnormalities such as these might occur coincidentally. That is, the abnormality might have nothing to do with disease (some parts of our genome do not contain critically important genes, and chromosome abnormalities affecting these regions might be found in a small percentage of normal people). Alternatively, the abnormality causes the disease by affecting the expression or structure of a gene or genes in that region. That would be more likely if the same DNA region were disrupted in two or more unrelated individuals with the same disease, or if the abnormality occurred *de novo* in a sporadic case (the affected individual has no family history of the disorder, and the abnormality is not present in either parent).

Chromosomal abnormalities occur rarely, so this approach can never be a general one for identifying disease genes. But it has been a useful for some disorders, notably dominantly inherited disorders with a severe congenital phenotype (because they normally cannot reproduce, affected individuals occur as sporadic cases, making linkage analyses problematic). Metaphase or prometaphase chromosome preparations can be prepared readily from blood lymphocytes, and then stained with DNA-binding dyes to reveal altered chromosome banding patterns that indicate a chromosome abnormality such as a translocation, deletion, inversion, and so on (Box 7.2 on p 204–5 gives the chromosome banding methodology).

Balanced translocations and inversions can be particularly helpful because, unlike large deletions, they may involve no net loss of DNA, and the underlying disease gene might be expected to be located at, or close to, a breakpoint that can readily be identified. See **Table 8.1** for some examples.

TABLE 8.1 EXAMPLES OF SUCCESSFUL GENE IDENTIFICATION PROMPTED BYTHE IDENTIFICATION OF DISEASE-ASSOCIATED CHROMOSOMAL ABNORMALITIES

Disorder	Chromosome abnormality	Comments	PMID
Duchenne muscular dystrophy	an affected boy with a cytogenetically visible deletion at Xp21.3 and a woman with a balanced Xp21; 21 p12 translocation	positional cloning strategies identified genes within the deletion/translocation breakpoint, finally implicating the giant dystrophin gene	2993910; 3001530
Sotos syndrome	a girl with a *denovo* balanced translocation with breakpoints at 5q35 and 8q24.1	the disease gene, *NSD1,* was found to be severed by the 5q35 breakpoint	11896389

PMID, PubMed identifier of relevant publications, at http://www.ncbi.nlm.nih.gov/pubmed/—see glossary.

Genome-wide screens for large-scale duplications and deletions are also available through different modern methodologies involving DNA hybridization or DNA sequencing

approaches. We will describe these methods in detail within the context of genetic testing in Section 11.2.

Exome sequencing: let's not bother getting a position for disease genes!

Although many genes underlying monogenic disorders have been identified by the methods used above, some monogenic disorders have not been well studied because they are very rare, or because they are not readily identifiable. There may be difficulties, for example, in recognizing individual phenotypes within complex sets of overlapping phenotypes, such as intellectual disability. More than 7000 monogenic disorders are estimated to exist, and although the methods above have been very successful in gene identification, a new approach was needed to identify genes in the substantial proportion of monogenic disorders that are very rare or where there is difficulty in distinguishing the phenotype from related disorders.

The new approach? That would be massively parallel DNA sequencing (also called next-generation sequencing), which we introduced in Section 3.3; details of two popular methods are given in Chapter 11). This new approach offers a vast increase in sequencing capacity, and the time taken to sequence a human genome has fallen from several years to a few days, or even a few hours. And the expense has dropped drastically. The inevitable consequence has been an explosion of genome sequencing.

The ability to sequence whole genomes (and therefore all genes) both rapidly and cheaply means that disease genes can often be identified without any need to first find a chromosomal position for the disease gene. Because the vast majority of disease genes are protein-coding genes and given that the great majority (perhaps 85 % or so) of currently known disease-causing mutations occur in the exons of protein-coding genes, whole-genome sequencing initially appeared a rather laborious and costly approach to identify a disease gene. However, all the exons of the protein-coding genes together account for only just over 30 Mb of our DNA; sequencing this fraction, just over 1 % of the genome, appeared an easier and cheaper option than genome sequencing.

Exome is the collective term for all exons in the genome. Operationally, exome sequencing has largely focused on the exons of protein-coding genes (RNA genes have been a low priority, mostly because they are viewed as very infrequent direct contributors to disease). Exome sequencing involves first capturing exons from the DNA of affected individuals, and then sequencing the captured DNA. In practice, exome capture is designed to capture exons with a little flanking intron sequence (to cover splice junctions) plus DNA sequences specifying some miRNAs; hybridization with a control set of cloned exon sequences allows capture of the desired exons, as shown in **Figure 8.7**.

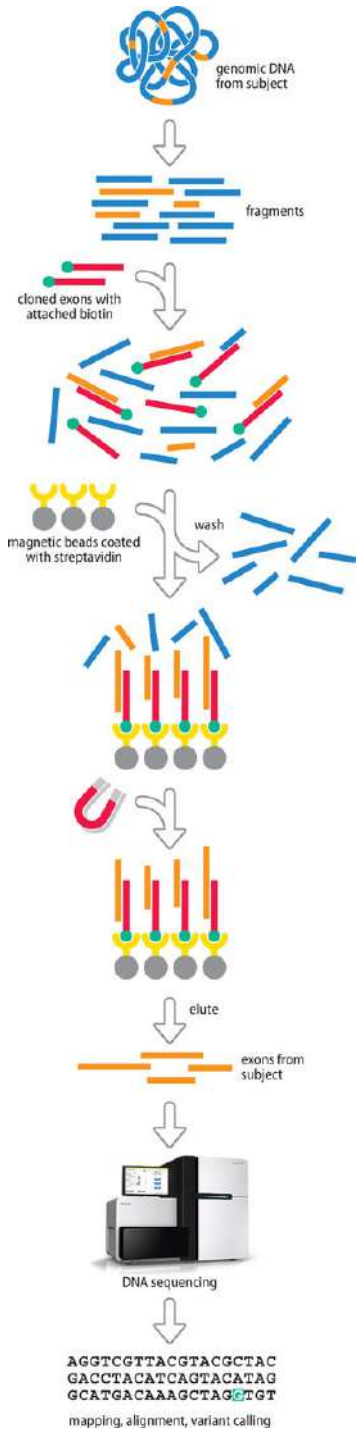


Figure 8.7 Exome capture and sequencing. A genomic DNA sample to be analyzed is randomly sheared and the fragments are used to construct a DNA library (library fragments are flanked by adapter oligonucleotide sequences; not shown). The DNA library is then enriched for exon sequences (orange rectangles) by hybridization to DNA or RNA baits that have been designed to represent human exon sequences (red rectangles). The baits have a biotin group (green circle) attached to their end. After capturing exon sequences from the test sample DNA by hybridization, the biotinylated baits can be selected by binding to magnetic

beads (gray) coated with streptavidin (yellow goblet shapes)—streptavidin has an extremely high affinity for binding to biotin, and the streptatividin–biotin–exon complexes can be removed using a magnet. The captured exon sequences from the sample DNA are subjected to massively parallel DNA sequencing and the data are interpreted as described in the text. (Adapted from Bamshad MJ et al. [2011] *Nat Rev Genet* 12:745–755; PMID 21946919. With permission from Macmillan Publishers Ltd.)

To identify a gene for a rare disorder, sequencing of the exomes from just a few affected individuals is often sufficient. That is so because clearly deleterious mutations (frameshifting and nonsense mutations, and some types of nonconservative amino acid substitution) can often be easily identified in protein-coding sequences. Because we have some non-essential genes, each of us carries a surprising number of deleterious mutations like this, which are scattered throughout the genome and vary from individual to individual. However, unrelated individuals with the same single-gene disorder might be expected often to have causative mutations in the same gene. Where there is parental consanguinity, exome sequencing can sometimes be used to identify genes underlying an autosomal recessive condition by exome sequencing of affected members of a single family.

Since its first successful application in identifying disease genes in 2009, exome sequencing has been dramatically successful in identifying genes underlying very rare autosomal recessive and congenital dominant disorders (neither of which is amenable to linkage analyses because of a lack of suitable families); see **Table 8.2** for examples. It has also been important in the case of extremely heterogeneous phenotypes—in one early study, 50 novel genes were identified for recessive cognitive disorders (**Figure 8.8**). As sequencing costs drop, future studies may use whole genome sequencing.

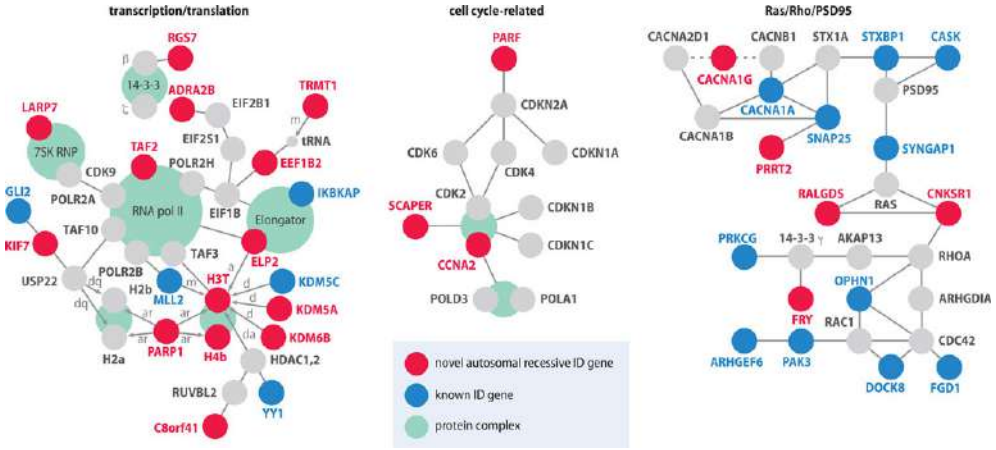


Figure 8.8 Known and novel genes for intellectual disability are part of protein and regulatory networks. In a single exome sequencing study, 50 novel genes were identified for recessive intellectual disability (ID). The novel genes were predicted to encode components of protein and regulatory networks that had been implicated by studies of known genes for intellectual disability, including the transcriptional/translational network, the cell cycle-related network, and the Ras/Rho/PSD95 network.

(Adapted from Najmabadi H et al. [2011] *Nature* 478:57–63; PMID 21937992. With permission from Macmillan Publishers Ltd.)

TABLE 8.2 EXAMPLES OF SUCCESSFUL USE OF EXOME SEQUENCING TO IDENTIFY GENES CAUSING MONOGENIC DISORDERS

Disease	Type of disorder	Origin of exomes	Mode of inheritance	OMIM number	Gene locus	PMID
Miller syndrome	congenital disorders of development	mostly unrelated sporadic cases	AR	263750	*DHODH*	19915526
Kabuki syndrome type1			AD	147920	*KMT2D (MLL2)*	20711175
Schinzel-Giedion syndrome			AD	269150	*SETBP1*	20436468
Osteogenesis imperfect type VI	connective tissue disorder	affected sibs born to consanguineous parents	AR	613982	*SERPINF1*	21353196
Spastic paraplegia 30	early-onset neuromuscular disorder		AR	610357	*KIF1A*	21487076

For a review of different strategies for exome sequencing to identify disease genes, and of successful applications, see Gilissen C et al. (2012) *Eur J Hum Genet* 20:490–497; PMID 22258526. AD, autosomal dominant; AR autosomal recessive. (PMID, PubMed identifier at http://www.ncbi.nlm.nih.gov/pubmed/ as explained in the glossary.)

8.2 APPROACHES TO MAPPING AND IDENTIFYING GENETIC SUSCEPTIBILITY TO COMPLEX DISEASE

As outlined in Section 8.1, various strategies, often initially using gene linkage analyses, have led to identification of numerous causative genes for Mendelian disorders. But Mendelian disorders are rare; complex common diseases that are so much more difficult to analyze represent the great bulk of genetic disease. Family-based genetic linkage analyses have rarely been successful in these cases. Instead, there has been heavy reliance on mapping genetic susceptibility to complex disease by *association* analyses, notably genomewide association (GWA) studies. In this section we explain the background to these endeavors, the progress made, and the limitations of GWA studies. But first we provide some background on the complexity of common genetic disorders.

The polygenic and multifactorial nature of common genetic disorders

Investigations into monogenic disorders became a principal focus in medical genetics after the DNA cloning and sequencing revolutions began in the 1970s. We are easily accustomed to the idea of disease arising from causative mutation at a single gene locus. And up until now the kinds of genetic disease that we have described are single-gene disorders, mostly of neonatal or childhood onset, where genetic variants at a single gene locus are sufficient to cause disease, either in a nuclear gene (by Mendelian inheritance) or in a mitochondrial gene (by maternal inheritance).

To demonstrate Mendelian inheritance, a character must depend almost exclusively on the genotype at a single locus, virtually regardless of genotypes at every other locus and of the environment, history and lifestyle of the person bearing the trait. Monogenic disorders are inevitably therefore the rare exceptions and the way in which we view the inheritance mechanism for monogenic disorders is quite different for the great bulk of genetic disease.

The two quite different ways of viewing inheritance mechanisms

In the second half of the nineteenth century, two pioneering researchers independently laid down the foundations of modern genetics. Gregor Mendel focused on the independent combination of *discrete* characters (such as the famous yellow and green peas), while Francis Galton, Darwin's second cousin, used statistical principles to describe *continuous* phenotypes (height, weight, and so on).

Up until the present time the distinction of discrete versus continuous characters continues to reflect a fundamental dichotomy in the way we view mechanisms of inheritance. The dichotomy is somewhat reminiscent of the particle-wave duality of quantum physics, but unlike in quantum physics, the basic theory to reconcile the different Mendelian and Galtonian views of the mechanisms of inheritance was worked out relatively quickly. In the early twentieth century Ronald Fisher formulated the infinitesimal (= polygenic) model of inheritance. In this model, Fisher envisaged that continuous features were determined by an almost infinite number of Mendelian factors, and that the large variance among children of the same parents was due to the segregation of those factors that were heterozygous in parents.

Given the complex interactions between gene products in the biological pathways in our cells and the diverse interactions with environmental components, single-gene disorders simply have to be the rare exceptions. Common genetic disorders—such as type 1 and 2 diabetes, coronary artery disease, stroke, rheumatoid arthritis, Alzheimer disease, and so on—are *polygenic*. That is, genetic variants at multiple loci are important contributors to the phenotype. And environmental factors also may make very significant contributions to the pheno-type, such as in the case of cigarette smoking for a range of common genetic diseases. Because of the genetic contributions of multiple genes plus environmental factors, common genetic diseases are said to be **multifactorial**.

The not quite clear-cut division between single-gene and multifactorial disorders

In reality, the division between single-gene disorders and multifactorial genetic disorders is not quite so clear cut. In virtually no case is the phenotype of a single-gene disorder entirely attributable to a single gene locus (Figure 7.21 gives the example of b-thalassemia). We can see the effects of other loci when the phenotype in a "single-gene" disorder differs between affected members of the same family (as in Figure 5.14). And when single-gene knockouts are made in mice, the phenotype can vary very significantly in different strains of mice.

The difference in phenotype between affected family members with a single gene disorder, who would be expected to have the same disease allele(s), or in different strains of mice having identical gene knockouts, is due to having different alleles at one or more **modifier** loci. The product of a modifier gene interacts with the disease allele in some way: it may regulate expression of the disease allele, for example, or it may interact in the same pathway as the product of the disease allele, affecting its function (see Section 7.9 for examples of modifier genes).

What distinguishes a single-gene disorder is that, although there may be minor effects from variants at other genetic loci, rare genetic variants at a primary gene locus have a very great effect on the phenotype. By contrast, it is usual to think of a polygenic disorder as being one in which the genetic susceptibility to disease risk is not dominated by a primary locus where individual variants can have extremely strong effects on the phenotype.

There may be a predominant genetic locus in a true polygenic disease, but its effect would not normally be large: variants at that locus would not be expected to be necessary or sufficient to cause disease. However, as described later, quite often problems arise out of heterogeneity: individual complex diseases can be a mix of related phenotypes, and in some of these diseases rare variants can be of quite strong effect.

The phenotypes of polygenic diseases are also influenced by nongenetic factors, including environmental factors (often working though epigenetic modification) and also chance (*stochastic*) factors. Environmental factors can sometimes have a strong influence in certain monogenic disorders, but their effects are very important right across the spectrum of polygenic disorders.

Because a combination of multiple genetic susceptibility loci and environmental factors is involved, common genetic diseases have been considered to be complex or multifactorial diseases. Underlying polygenic theory is the idea that there is a *continuous susceptibility to the disease within the population,* and that disease manifests when a certain threshold is exceeded (Box 8.2).

BOX 8.2 POLYGENIC THEORY AND THE LIABILITY THRESHOLD CONCEPT TO EXPLAIN DICHOTOMOUS TRAITS

Human traits can be divided into two classes. Some, like diseases, are *dichotomous* (you either have the trait or you do not). Others, such as height or blood pressure, are *continuous* (or *quantitative*) traits—everybody has the trait, but to differing degrees. To explain quantitative traits, polygenic theory envisages the additive contributions of variable alleles from multiple loci. Many alleles at the underlying **quantitative trait loci** (**QTL**) might have subtle differences (causing modest changes in expression levels of certain genes, for example); different combinations of alleles at multiple loci might then produce continuous traits (adult height, for example, is known to be governed by genetic variants at a minimum of 180 loci).

Polygenic theory can be extended to also explain dichotomous traits using the concept of a *liability threshold*. The idea is that there is a continuous liability (susceptibility) to disease in the population, but only people whose susceptibility exceeds a certain threshold will develop disease. For each complex disease, the susceptibility curve for the general population will be a normal (bell-shaped) curve: most individuals will have a medium susceptibility, with smaller numbers of people having low to very low susceptibility or high to very high susceptibility (**Figure 1A**). Only a small percentage of the general population will have a susceptibility that exceeds the threshold so that they are affected by the disease (shown in red in Figure 1).

Close relatives of a person affected by a complex disease have an increased susceptibility to that disease (Figure 1B). The affected person must have a combination of different high-susceptibility genes; relatives will share a proportion of genes with the affected person and will therefore have an increased chance of having high-susceptibility genes.

By chance, some relatives may have only a few of the high-susceptibility genes, but others may have many high-susceptibility genes in common with the affected person. While disease susceptibility can show wide variation among first-degree relatives (parent and child, sibs), the overall effect is that the curve—and the median susceptibility (dashed vertical line)—is displaced to higher susceptibilities to the disease (Figure 1B). Because the threshold remains the same, more individuals are affected, and the relatives that are most closely related to an affected person are more likely to be affected.

VARIABILITY IN THE LIABILITY THRESHOLD FOR AN INDIVIDUAL DISEASE

Thresholds of susceptibility to a complex disease are not absolutely fixed—they can show differences between the two sexes, for example. For many auto-immune disorders, women have significantly more disease risk than men. The reverse is true for some other conditions. For example, pyloric stenosis occurs in about 1 in 200 newborn males, but only in about 1 in 1000 newborn females. That is, a double threshold exists, one for females and one for males. The female threshold for pyloric stenosis is farther from the mean than that

for the male. However, because it takes more deleterious genes to create an affected female, she has more genes to pass on to the next generation. Her male offspring are at a relatively high risk of being affected when compared with the population risk.

The threshold model accommodates environmental effects by postulating that such effects can reposition the threshold with respect to the genetic susceptibility. Protective environmental factors move the threshold to higher genetic susceptibilities; other environmental factors can increase the risk of disease by moving the threshold to lower genetic susceptibilities.

Complexities in disease risk prediction

For a common genetic disease, no single variant causes the disease. Instead, variants at multiple loci can act independently to increase or decrease the risk that a person will develop the disease. And non-genetic factors—environmental factors, lifestyle choices, and sometimes even chance—can make very signifi-cant contributions to the risk of disease.

For the reasons above, common genetic diseases do not give the typical dominant and recessive patterns seen in Mendelian diseases. Sometimes, however, a common genetic disease can show some evidence of running in families. A minority of families may have multiple affected members, and although rare, some families seem to show Mendelian or quasi-Mendelian transmission (such as in early-onset forms of Alzheimer disease, diabetes, Parkinson disease, and various types of cancer). Frequently, however, there is little evidence of family history; quite commonly an affected individual appears as a sporadic case.

Relatives share genes, and so may share variants predisposing to a common disease and may have a higher risk of developing the condition than the average risk across a population. The **relative risk** of developing a common genetic disease (also called the **risk ratio**) is the disease risk for a relative of an affected person divided by the disease risk for an unrelated person in the general population and is denoted by the symbol l. Different values for l can be calculated for different degrees of relationship, such as l_s which expresses the relative risk for sibs (brother or sisters) of an affected person. Table 8.3 gives illustrative values of l_s for certain monogenic and multifactorial conditions.

TABLE 8.3 EXAMPLES SHOWING CONTRASTING RELATIVE RISKS AND LIFETIME RISKS FOR MENDELIAN AND MULTIFACTORIAL DISORDERS

Disease	λ_s (relative risk)	Lifetime risk (to age 80)

Disease	λs (relative risk)	Lifetime risk (to age 80)
Huntington disease	5000	0.01 %
Cystic fibrosis	500	0.05%
Alzheimer disease (l.o.)*	3	17%
Breast cancer, female	2	12%
Type 1 diabetes	15	1%
Type 2 diabetes	3	15%

Monogenic disorders are very rare. But if a person has an affected sibling their relative risk (λs) is very high, such as 5000 for Huntington disease (the 50% risk for a dominant disease is 5000 times greater than the incidence of 0.01 % in the general population). The lifetime risk for common diseases such as Alzheimer disease or breast cancer is high, and the relative risk is low.

* l.o. = late onset

The calculation of disease risk for complex diseases is therefore often quite different than for Mendelian disorders. In the latter case, disease risks are mostly based on theoretical calculations that remain quite stable. If a child is born to a couple with an autosomal recessive condition, for example, then one can normally assume that both parents are carriers and that all subsequent pregnancies carry a 1:4 risk of an affected fetus. (For some disorders, however, there can sometimes be complications, often because of low penetrance and variable expressivity.)

For some complex diseases, we are beginning to accumulate information on major predisposing genetic and environmental factors; this may ultimately lead to a more informed measure of disease risk than has previously been possible. Traditionally, a lack of knowledge about the predisposing risk factors has meant that for complex diseases the disease risk has often been *empiric:* that is, it is based on observed outcomes in surveys of families and is often *modi-fied according to past incidence of disease*. For example, a couple who have had one baby with a neural tube defect would be quoted a specific recurrence risk for a subsequent pregnancy that would be based on observed frequencies in the population; if they do go on to have a second affected child, the disease risk for a subsequent pregnancy would now be substantially higher. The real disease risk would not have changed; the birth of the second affected child helps us to recognize that the parents carry more high-susceptibility genes than was apparent after the birth of their first affected child.

Difficulties with lack of penetrance and phenotype classification in complex disease

Researchers seeking to identify the genetic susceptibility in multifactorial disorders are confronted by multiple challenges. In the sections below we cover some of the technological difficulties that had to be surmounted. Here we consider more intrinsic difficulties arising

from the general lack of penetrance, or from problems with defining and classifying disease phenotypes.

Recall that reduced penetrance can be a feature of some monogenic disorders, such as imprinting disorders. But, in general, DNA variants at a Mendelian disease locus are of very strong effect and are therefore highly penetrant. As a result, affected people typically have a disease-associated genotype; unaffected people do not. For complex disease, however, the picture is different. If we discount rare Mendelian subsets of complex disease, reduced penetrance is the norm, simply because multiple genes make small contributions to the phenotype. That is, a DNA variant strongly associated with a complex disease is often at best a **susceptibility factor**: its overall frequency should be significantly increased in affected individuals compared with controls, but normal people can quite often possess the variant conferring disease susceptibility and affected people can quite often lack it. There is a clear contrast, therefore, between studying the genetics of monogenic disorders and multifactorial diseases: in the former we look for DNA variants that are fully responsible for causing the disease, and in the latter, we seek genetic susceptibility factors that predispose to disease.

Phenotype classification and phenocopies

In many complex diseases the phenotype can have many variable components. As detailed in Section 5.3, monogenic disorders can also have variable pheno-types, but here the high penetrance and the range of phenotypic features in multiple affected individuals reported in the literature allow us to be clear about which aspects of a person's phenotype are disease components, and which are not. In complex diseases, however, the situation is not so simple, and phenotype delineation and classification can be a major problem.

Some affected people, who do not have genotypes commonly associated with the disease, are **phenocopies** that have been wrongly classified as having the disease under study. For some phenocopies the disease is caused by different genetic factors than expected. For example, accurate diagnosis of Alzheimer disease has traditionally relied on *post-mortem* brain pathology. If we wish to study living patients, various clinical tests can be conducted, but a subsequent diagnosis of Alzheimer disease is often provisional ("*probably* Alzheimer disease"); post-mortem examination might subsequently show a different type of dementia, such as Lewy dementia, frontotemporal dementia, and so on. For other phenocopies, the phenotype might have an environmental origin.

According to the condition, defining the disease phenotype can be straightforward or very challenging. At one extreme are conditions in which there is a very recognizable and rather specific phenotype. For example, in primary biliary cirrhosis, the most common autoimmune form of chronic liver disease, affected individuals have markedly similar phenotypes, and have signature autoanti-bodies that are specifically directed against the E2 subunit of the

mitochondrial pyruvate dehydrogenase complex. At another extreme are some behavioral and psychiatric conditions, in which even classifying individuals as affected and unaffected can sometimes not be straightforward. The pathology might not be well defined and there can be heavy reliance on interviews (and subjective information). Here, clear diagnostic criteria are of paramount importance.

Deciding which aspects of a person's phenotype are components of a complex disease is not easy, and deciding how far we should lump together different, but clearly related, phenotypes is a significant issue in complex disease studies. If two first cousins have had different types of congenital heart malformation, for example, should we consider the phenotypes as independent occurrences, or lump them together and report two affected individuals in one family?

Estimating heritability: the contribution made by genetic factors to the variance of complex diseases

Phenotypes are determined both by genetic factors and by nongenetic factors (often described as environmental factors, but also comprising stochastic factors). The *variance* (V) of a phenotype is a statistical term defined as the square of the standard deviation. The total variance of a phenotype, $V_P = V_G + V_E$ (where V_G is the genetic variance and V_E is the environmental variance).

The genetic variance, V_G, is the sum of three components: (i) additive genetic effects (the combined contributions from different loci); (ii) dominance effects (interaction of alleles at a single locus); and (iii) interaction effects (interaction between genes at different loci; the effect of genes at several loci may not be simply additive if they interact with each other).

The **heritability** (h^2) is that proportion of the variance that can be attributed to genetic factors—that is, $h^2 = V_G/V_P$. Possible values for h^2 range from 0 (no genetic factors involved) to 1 (exclusively due to genetic factors). As described below, the ratio of genetic to environmental involvement in the etiology of a disease varies according to the disease class.

Despite its importance in the prediction of disease risk, heritability has often been misunderstood. First, it must be appreciated that heritability is not a fixed property. Secondly, it describes a population, not an individual. More accurately, it describes the genetic contribution to variance *within a population and in a specific environment at a certain time*. To estimate the heritability of a complex trait or disease, the incidence of disease is compared in genetically related individuals. To do this, three types of study have been undertaken: family studies, twin studies, and adoption studies.

Family studies

Having a relative with a complex disease increases your risk of developing that disease. The **risk ratio (λ)** is defined as the disease risk to a relative of an affected person divided by the

disease risk to an unrelated person in the general population. The comparative disease risk for a sib (brother or sister) of an affected individual is designated as ls. Unlike a monogenic disorder (for which the risks to family members are fairly precisely defined, according to simple theoretical calculations), the risks for complex diseases have quite often been empiric; that is, they have often been based on surveys of disease incidence in families.

Because l is a measure of how important genetic factors are in the etiology of the disease, extremely high ls values are found in monogenic disorders (Table 8.3). For example, in populations of western European origin, the general lifetime risk of cystic fibrosis is about 1 in 2000, but the risk to a fetus is 1 in 4 if the parents have had a previously affected child. In that case, ls = 1/4 divided by 1/2000 = 500. For some complex diseases, ls values can be quite high, such as 25 for Crohn's disease, 20 for multiple sclerosis and 15 for type 1 diabetes.

There is, however, a significant drawback in using family studies to infer genetic factors: family members would normally be exposed to some common environmental factors as a result of the shared family environment.

Twin studies

Twin studies measure how often the twins are concordant (both have the disease) or discordant (only one is affected). Monozygotic twins are genetically identical but dizygotic twins, like any pair of sibs, share 50 % of their genes. Regardless of zygosity, twins should be exposed to rather similar environment factors.

TABLE 8.4 DEGREE OF CONCORDANCE BETWEEN TWIN PAIRS FOR SELECT COMPLEX DISEASES, AVERAGED FROM MANY STUDIES

Disease	Concordance(%)	
	In MZ twins	In DZ twins
Type 1 diabetes	42.9	7.4
Type 2 diabetes	34	16
Multiple sclerosis	25.3	5.4
Crohn's disease	37	10
Ulcerative colitis	7	3
Alzheimer disease	32.2	8.7
Parkinson disease	15.5	11.1
Schizophrenia	40.8	5.3

MZ, monozygotic; DZ, dizygotic.

Table 8.4 lists observed concordance rates for monozygotic and dizygotic twins for various complex diseases. There are two major points to note. First, there are significant discordance rates between monozygotic twins—genetics is not everything! Secondly, it is clear that in certain diseases the concordance between monozygotic twins is much greater than that between dizygotic twins. Diseases in which genetic factors have a large role show a relatively high concordance in monozygotic twins and a much lower concordance rate in dizygotic twins: the greater this ratio, the greater is the genetic contribution. Thus, for example, genetic factors would seem to be much more important in schizophrenia than in Parkinson disease.

The best type of study would seek to disentangle the contributions of genetic and environmental factors by tracking monozygotic twins separated at birth and raised in different environments, but this occurs so rarely that statistically valid sample sizes cannot be obtained.

Adoption studies

Adoption studies offer the best practical way of separating out the contributions made by genetic and environmental factors. Two types of study can achieve this. One type of study investigates children who were born to affected individuals but were then adopted at birth into an unaffected family (the children might be expected to share genetic factors with their biological parents but be exposed to different environments).

A second type of adoption study starts with adopted people who suffer from a specific common genetic disease that can run in families. The object is to seek evidence for the disease running in the biological family or their adoptive family. A celebrated example is a Danish study of 14 427 adopted people aged 20–40 years, 47 of whom were diagnosed with chronic schizophrenia. The 47 schizophrenia cases were extensively matched (for age, sex, social status of adoptive family, and number of years in institutional care before adoption) with 47 control non-schizophrenic adoptees from the 14 427 adoptees. The results showed quite clearly that genes rather than the family environment increased the risk for the offspring. Interested readers can find the data in the publication by Kety SS et al. (1994) *Arch Gen Psychiatr* 51:442–455; PMID 8192547.

Adoption studies are the gold standard for teasing out the separate contributions of genetic factors and environmental factors, but they are difficult to carry out. Often, little is often known about the biological family, and intrusive approaches to the biological family may be undesirable. Efficient adoption registers may be lacking in many countries, and, in the interests of the child, adoption registers tend to choose a foster family resembling the biological family. Thus far, adoption studies have largely been carried out for psychiatric conditions (where there has been great interest in the contributions of nature and nurture to disease development).

Variation in the genetic contribution to disorders

Heritability studies have indicated that genetic factors make different contributions to different diseases. Monogenic disorders are primarily determined by genetic factors (but in some cases environmental factors can make very significant contributions). By contrast, infectious diseases are primarily determined by environmental factors (exposure to the infectious agent). Here, however, host genetic factors also make a contribution, so that people vary in their susceptibility to an infectious disease and a small number of individuals may be disease-resistant.

For the great bulk of other complex diseases, genetic and environmental factors both make large contributions to the phenotype. In some complex diseases, such as schizophrenia, autism spectrum disorder, Alzheimer disease, type 1 diabetes, multiple sclerosis, and Crohn's disease, there is a strong genetic contribution; in others, such as Parkinson disease and type 2 diabetes, genetic factors seem to be less important.

The heritability of an individual disease varies from one population to another. It can also vary in the same population in response to a changing environment. Consider phenylketonuria. As detailed in Clinical Box 9 on page 234, a deficiency of phenylalanine hydroxylase produces elevated phenylalanine and toxic by-products that can result in cognitive disability. In the recent past the disease was almost wholly due to genetic factors, and so the heritability was extremely high. In modern times, neonatal screening programs in many countries allow early detection and treatment using low-phenylalanine diets. Now, in societies with advanced health care, phenylketonuria results mostly from environmental factors that lead to failure to deliver the treatment (inefficiency in health care systems, reluctance of families to seek out treatment, non-compliance with the diet, and so on).

The incidence of some complex diseases is very dependent on environments that can change very significantly with time. Thus, the huge recent increase in type 2 diabetes (mostly as a result of increasingly unhealthy diets and lack of exercise) means that the heritability of this disorder in many populations is now much reduced when compared with just a few decades ago.

The very limited success of linkage analyses in identifying genes underlying complex genetic diseases

The linkage analyses used to map genes for monogenic disorders are said to be *parametric*—the data can be analyzed only if a specific genetic model is assumed. The model needs to give details of certain key parameters: the mode of inheritance, disease gene frequency, and penetrance of disease genotypes. Parametric linkage analyses have been very successful in mapping genes for Mendelian disorders, but they have had very limited applicability in complex disease (because of the general difficulty in providing all the required parameters). They have, however, been of use when applied in two ways: (i) in analysing rare Mendelian

subsets of complex diseases; (ii) when using certain non-parametric linkage analyses, usually to study affected sibs from large numbers of families.

Parametric linkage analyses in Mendelian subsets

Parametric linkage analyses are more readily applied when a complex disease shows very strong familial clustering. A near-Mendelian pattern may simply reflect the chance occurrence of several affected people in one family. If, by chance, most members of a family possess multiple genetic variants conferring high susceptibility to disease (not necessarily at the same loci in different affected family members), a single common disease-susceptibility allele (that has been transmitted in a Mendelian fashion within that family) might be enough to tip the balance past the susceptibility threshold.

For several complex diseases there are also subsets in which the disorder shows clear Mendelian inheritance. That is most obvious when there is dominant inheritance. In early-onset Alzheimer disease, for example, some large pedigrees, such as the one shown in **Figure 8.9**, permitted identification of three disease genes: the amyloid precursor protein gene (*APP*), and two presenilin genes involved in processing the APP protein. In some other complex diseases, such as Parkinson disease, a number of autosomal recessive pedigrees have led to gene identification. We detail the genes identified through Mendelian subsets of Alzheimer and Parkinson disease in Section 8.3.

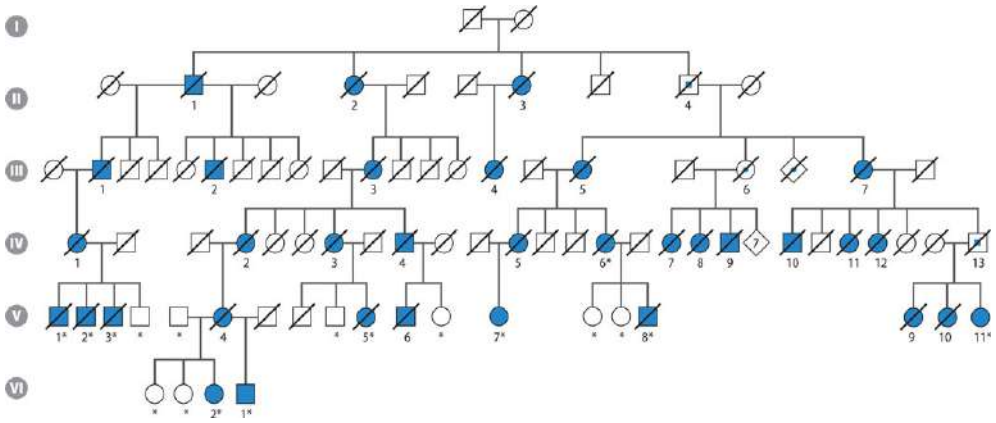


Figure 8.9 An exceptional pedigree showing dominantly inherited Alzheimer disease. Affected members of this pedigree had early-onset Alzheimer disease (average age at onset 46 ± 3.5 years). They were subsequently shown to have a mutation in the presenilin 1 gene, *PSEN1*. (From Campion D et al. [1995] *Neurology* 45:80–85; PMID 7824141. With permission from Wolters Kluwer Health.)

How might the phenotypes of a complex disease and one that segregates like a Mendelian disorder be so similar? The gene mutated in the Mendelian subset might also be a disease-susceptibility locus for the complex disease (with a rare, highly penetrant variant in the Mendelian subset, and a common variant of weak effect at the same locus in the complex

disease). Or, if genes mutated in Mendelian subsets are not significant disease-susceptibility loci for a complex disease, the common pathogenesis might suggest that at least different genes associated with the two forms of the disease are part of a common biological pathway or process.

Affected sib-pair and other nonparametric linkage analyses

Nonparametric methods of linkage analysis do not require any genetic model to be stipulated, and so can generally be applied to analyzing complex disease. They rely on the principle that, regardless of the mode of inheritance, affected individuals in the same family would tend to share not just major disease-susceptibility genes but also the immediate chromosomal regions. That is, a major disease-susceptibility locus and a very closely linked marker would show a strong tendency to be co-inherited within affected individuals in the same family (because of the very low chance of recombination between the marker locus and the disease locus).

Nonparametric linkage studies occasionally use samples from all affected family members, but it is usually more convenient to simply use affected sib pairs. The aim here is to obtain genome-wide marker data in affected sibs from multiple families and then identify chromosomal regions that have been shared more often than would be predicted by random Mendelian segregation. As sibs share 50 % of their genes, affected sibs need to be studied in many families with the same complex disease. For marker loci that are not linked to a major disease susceptibility gene, sibs would be expected to share 50 % of alleles on average (some sib pairs might share 2, 1, or 0 alleles by chance, but the *overall average* across all sets of sibs would be 1 allele in common). For marker loci close to a major disease susceptibility gene, affected sibs would be expected to share significantly more than 50 % of alleles (see **Figure 8.10** for the principle).

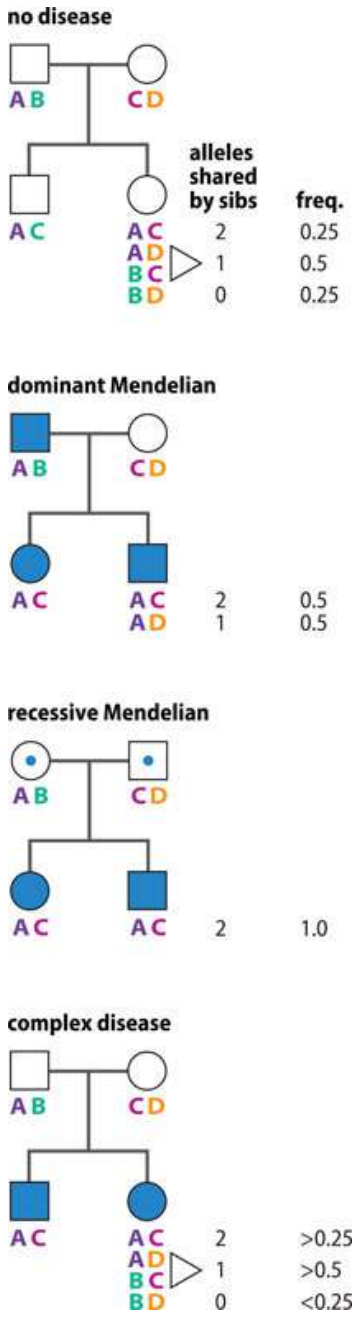


Figure 8.10 Principle of affected sib-pair analysis. By random segregation, any pair of sibs share 2, 1, or 0 parental haplotypes in the relative proportions 1:2:1. Pairs of sibs who are both affected by a dominant Mendelian condition must share the segment that carries the disease allele, and they may or may not (a 50:50 chance) share a haplotype from the unaffected parent. Pairs of sibs who are both affected by a recessive Mendelian condition necessarily share the same two parental haplotypes for the relevant chromosomal segment. For complex conditions, haplotype sharing greater than that expected to occur by chance may allow the identification of chromosomal segments containing susceptibility genes.

Affected sib-pair analyses are comparatively easy to carry out (they need samples from just a few people per family) and robust (the method makes few assumptions). But there are inevitable limitations. Because any individual susceptibility factor is neither necessary nor sufficient for a person to develop a complex disease, the underlying genetic hypothesis is weaker than for Mendelian conditions. This means that finding statistically significant evidence for a disease susceptibility factor is going to be harder.

The calculations in **Table 8.5** show that under ideal conditions affected sib-pair analyses can be carried out with reasonable numbers of samples (typical studies use a few hundred sib pairs). But if the effects are weak, unfeasibly large numbers of samples are needed to detect them, and the studies will be defeated if there is a high degree of heterogeneity (individual susceptibility factors will operate in just a small proportion of families).

TABLE 8.5 NUMBER OF AFFECTED SIB PAIRS NEEDED TO DETECT A DISEASE SUSCEPTIBILITY FACTOR

Relative risk of disease (γ)	**Probability of allele sharing by affected sibs** (γ)		**Number of affected sib pairs needed to detect the effect***	
	At $p = 0.1$	**At $p = 0.01$**	**At $p = 0.1$**	**At $p = 0.01$**
1.5	0.505	0.501	115481	7868358
2.0	0.518	0.502	9162	505272
2.5	0.536	0.505	2328	103007
3.0	0.556	0.509	952	33 780
4.0	0.597	0.520	313	7253
5.0	0.634	0.534	161	2529

Here p is the frequency of the disease susceptibility allele and q is the frequency of normal alleles at the disease susceptibility locus (sothat $p+q = 1$).The relative risk of disease (γ) is a measure of how the disease risk changes when comparing persons with the susceptibility factor to those without. The calculations are based on the formulae derived by Risch & Merikangas in their 1996 paper (see Further Reading).

* Really the number of affected sib pairs with 80 % power to detect the effect. The take-home message is that unless the disease susceptibility factor is both quite common and confers a high disease risk, very many affected sib pairs are required to detect it.

Genome-wide nonparametric linkage scans require higher thresholds for statistical significance. Lod scores above 5.4 are considered highly significant evidence for linkage; scores between 3.6 and 5.4 are significant; and scores from 2.2 to 3.6 are suggestive. In practice, affected sib-pair analyses deliver typically modest lod scores that often do not reach statistically significant thresholds.

The review by Altmüller et al. (2001) under Further Reading reports an analysis of different nonparametric linkage studies for many diseases and underscores the difficulties. Of 10 genome-wide linkage studies of schizophrenia analyzed in that review, four were

unable to find any evidence of linkage, five found only suggestive evidence of linkage (lod scores between 2.2 and 3.5, but at many different regions on eight different chromosomes), and only one recorded more significant evidence of linkage. In addition to the lack of any great consistency in the results, getting independent replication of significant results proved very difficult. But in some cases, such as an 8p21 location, the initial finding was replicated in multiple populations. That finding eventually led to investigation of alleles of the neuregulin gene, *NRG1,* as a risk factor (the gene had been mapped to 8p21 and was known to be involved in synaptic transmission).

Identifying the disease-susceptibility gene

Even if a significant candidate chromosome region can be identified for a complex disease susceptibility locus, finding the implicated gene is often problematic in the absence of clues that suggest a candidate gene. That is so because sibs share large chromosome segments, and so candidate chromosome regions are generally very large (in Mendelian disorders, by contrast, the candidate chromosomal region can be progressively reduced by looking for rare recombinants between marker loci and the disease locus). To get closer to the disease susceptibility gene additional *linkage disequilibrium* mapping methods have sometimes been used. The same methods are regularly used in association analyses, and we will consider these in detail in the next section.

Despite the above difficulties, genome-wide linkage studies have had a measure of success in mapping susceptibility genes in complex disease to specific candidate regions that would then allow subsequent gene identification using other approaches. In addition to the schizophrenia-associated *NRG1* allele mentioned above, successes include mapping of genes conferring susceptibility to age-related macular degeneration to the 1q32 region, and genes conferring susceptibility to Crohn's disease at the 16q11-l6q12 region. These advances allowed subsequent *association analyses* to be targeted to these regions as described below, ultimately allowing identification of the *CFH* (complement factor H) gene at 1q32 and the *NOD2* gene at 16q12 as novel disease-susceptibility factors.

The *CFH* gene had previously been well known, but the gene that came to be known as *NOD2* was identified only very shortly before being implicated in Crohn's disease; it would provide the first molecular insights into the pathogenesis of this disease. In Crohn's disease an abnormal immune response is directed against various *nonself* antigens in the gut, including harmless (and often beneficial) commensal bacteria; the resulting accumulation of white blood cells in the lining of the intestines produces chronic inflammation.

The *NOD2* gene was finally implicated in Crohn's disease by identifying three comparatively common variants: two missense mutations and, notably, a frame-shift mutation that occurs near the end of the coding sequence and has a weak effect (**Figure 8.11**). In one survey 50% of 453 European patients were reported to have presumptive

pathogenic mutations in the *NOD2* gene. The three common mutations accounted for 81 % of the mutations; homozygotes or compound heterozygotes for these mutations are not uncommon in Crohn's disease, but are very rare in the normal population. A heterogeneous set of rare missense mutations were suggested to account for the remaining causal variants.

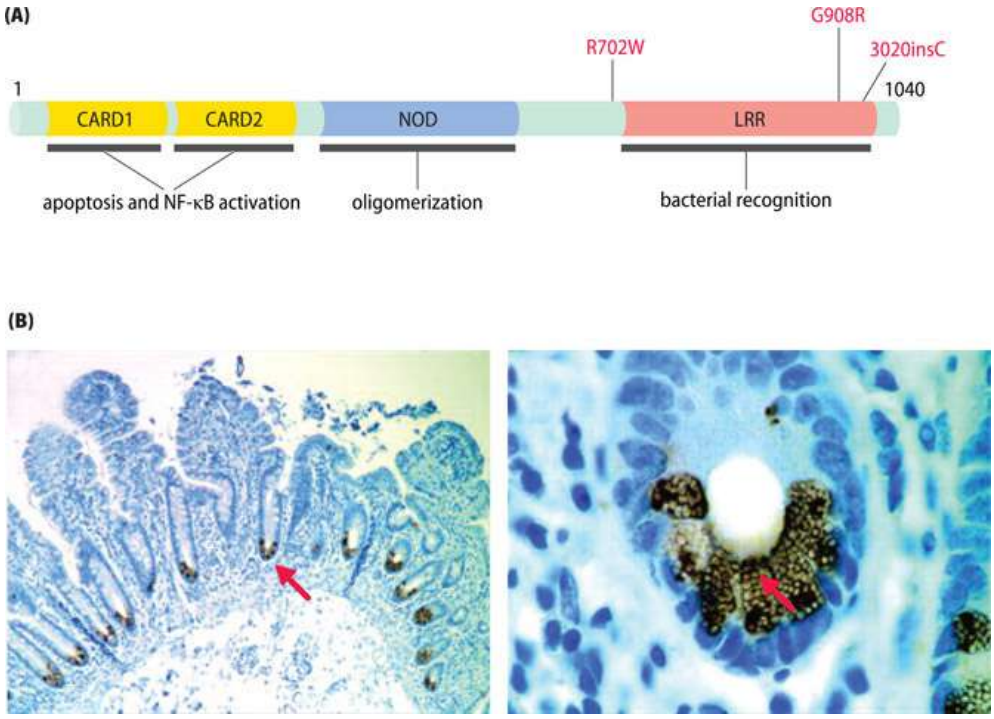


Figure 8.11 The Crohn's disease susceptibility factor NOD2: common variants and expression in Paneth cells. (A) Domain structure of the 1040-residue NOD2 protein and corresponding location of common variants associated with Crohn's disease (in red). The 3020insC variant appears to be a mild frameshift mutation; it inserts a cytosine, causing a stop codon to be introduced at the next codon position (codon 1008), eliminating just the final 33 amino acids. Like 3020insC, the missense mutations G908R and R702W are located within or close to the LRR domain. Domains: CARD1, CARD2, caspase-activating recruitment domains; NOD, nucleotide-binding oligomerization domain; LRR, leucine-rich repeats domain. The LRR domain is now known to bind to specific breakdown products of peptidoglycan, a major component of bacterial cell walls. (B) The NOD2 protein is predominantly expressed in Paneth cells, specialized secretory epithelial cells found at the base of intestinal crypts (arrows show examples of staining with a specific anti-NOD2 antibody). Paneth cells secrete certain anti-microbial peptides, notably a-defensins. (B, From Ogura Y et al. [2003] *Gut* 52:1591–1597; PMID 14570728. With permission from the BMJ Publishing Group Ltd.)

The NOD2 protein is now known to be part of the **innate immune system** (which produces the initial non-specific immune responses against pathogens). It has a C-terminal domain that recognizes a specific peptide motif found in a wide variety of bacterial proteins (see Figure 8.11). The three common DNA variants in Figure 8.11A seem to be partial loss-of-function mutations that impair the ability of NOD2 to recognize bacterial protein.

Gut flora (*microbiota*) include many microbes that are of active benefit to us. They help us derive additional energy through the fermentation of undigested carbohydrates, help us break down xenobiotics, and synthesize vitamins (B and K) for us. Although they are foreign microorganisms they are therefore tolerated by suppressing the usual innate immune responses. NOD2 works in this area by down-regulating those innate immune responses that require the Toll-like receptors; when NOD2 is impaired, a strong immune response is launched in response to the gut flora, causing inflammation.

The fundamentals of allelic association and the importance of HLA-disease associations

Linkage analyses inherently have very limited power to detect susceptibility factors in complex disease. They have been profitably used in rare Mendelian subsets to detect susceptibility factors with sizeable effects, as in the case of early-onset Alzheimer disease (the amyloid precursor protein gene A*PP* and the presenilin genes *PSEN1, PSEN2*); breast cancer (*BRCA1, BRCA2*); colon cancer (*APC*), and maturity onset type 2 diabetes (*GSK*, the glucokinase gene). When the alternative of affected sib pair analyses have been employed to get to a subchromosomal location, allelic association studies have often then been used to get to the gene (such as in helping to identify *NOD2* variants as susceptibility factors for Crohn's disease).

From the mid-2000s onwards, powerful genomewide association analyses supplanted linkage analyses as the route to identifying genetic factors in complex disease. We describe below how these are conducted, but first let us look into what exactly allelic association means.

The nature of allelic association

Whereas linkage is a *genetic* phenomenon, **association** is essentially a *statisti- cal* property that is simply concerned with unexpected frequencies for the cooccurrence of alleles (and/or phenotypes) in individuals within a population. **Figure 8.12** gives a summary of the essential differences between genetic linkage and allelic association.

LINKAGE		ASSOCIATION
a specifically **genetic** relationship	NATURE	a **statistical** relationship based on a frequency of co-occurrence that deviates from random expectation
a relationship between **loci**; needs comparison of **family** members	IN GENETICS	a **population**-based relationship between **alleles** or **phenotypes**
a marker locus and a disease locus are **linked** if they lie close together on chromosome, such as the HLA-B locus and the CYP21A2 (21-hydroxylase) locus	AN EXAMPLE	the HLA-B27 allele is **associated** with ankylosing spondylitis. In the UK >90% of patients exhibit it, but only about 8% of the general population

Figure 8.12 Genetic linkage and association compared.

In a population under study, if allele *A*1* at locus *A* is found to be significantly more frequent in people affected by a specific complex disease D than would be expected (from the individual population frequencies of *A*1* and the *D* gene), we would say that allele *A*1* is positively associated with disease, a disease-susceptibility allele. Conversely, if significantly less frequent in affected individuals, *A*1* would be a disease-resistance allele (negatively associated with disease).

Unlike linkage, the association of alleles in a population is not intrinsically genetic. There can be several ways in which the association can be explained, not all of which are genetic. Four possibilities are listed below:

- *Direct causation*. Simply by having allele *A*1*, a person is more susceptible to disease D. Somehow *A*1* confers an increased risk in the *population* of developing the disease (but *A*1* may not be necessary or sufficient for someone to develop the disease).
- *Linkage disequilibrium*. Allele *A*1* is not directly involved in pathogenesis but is nevertheless positively associated with disease. That can happen if allele *A*1* is located very close on the chromosome to a true susceptibility factor locus *B* where, for example, allele *B*2* is a high risk allele. Affected individuals would tend to have haplotype *A*1-B*2*; although both *A*1* and *B*2* would be positively associated with disease, allele *B*2* is the allele contributing to pathogenesis. We explain linkage disequilibrium more fully below.
- *Epistasis.* People who have disease D plus a high-risk allele *A*1* may be more likely to survive and have children if they also have allele *M*1* at *modifier* locus *M*. If so, *M*1* might also appear to be associated with disease. The modifier locus might make a product that interacts with other gene products in a pathogenetic pathway for disease D.
- *Population stratification.* A population happens to have various genetically distinct subpopulations and both disease D and allele *A*1* might just happen to be common in one subpopulation, whereupon allele *A*1* appears to be associated with disease. Eric Lander and Nicholas Schork gave the light-hearted example of association of the

*HLA*A1* allele and the "trait" of being able to eat with chopsticks (*HLA*A1* is more frequent in Chinese than in Caucasians).

Candidate gene association analyses, case-control studies, and odds ratios

Association studies have a long history, beginning long before the molecular genetics revolution, at a time when certain protein polymorphisms were commonly used, including alleles of the ABO and other blood group systems, and especially HLA polymorphisms. Because of their extremely high polymorphism, numerous alleles could be identified at each classical HLA gene locus using panels of antisera (that is, *serological* typing was performed instead of DNA typing).

In those early days, there was no possibility of carrying out genome-wide analyses to hunt for markers that showed association with a specific disease. That was so because unlike genetic linkage, which works over long ranges in DNA molecules, genetic association works over very short distances only. Genome-wide linkage analyses require just a few hundred markers distributed across the genome, but genomewide association analyses typically need many hundreds of thousands of markers to find a marker allele that is both associated with—and very tightly linked to—the disease-susceptibility allele.

Instead, in the absence of abundant DNA polymorphisms, there was simply the possibility of candidate gene studies: testing individual protein polymorphisms in turn to see if any showed significant evidence of association with a specific disease. Because of their key importance in T-cell function and cell-mediated immunity, HLA alleles were suspected to show disease associations and were comprehensively investigated using candidate gene association studies.

To investigate any disease associations, **case-control studies** are carried out in which genetic variants are typed in affected individuals (cases) and in controls from the population under study. In the smaller studies, such as the early HLA association studies, controls were selected to be individuals *known* to be unaffected, but in large-scale association analyses it is more convenient to use general population-based controls, for whom the disease status is simply unknown (with the proviso that the disease of interest is not too common in the population under study).

Different methods can be used to measure the disease risk for each tested genetic variant. One popular method is the **odds ratio**; that is, the odds of being affected when possessing a specific genetic variant divided by the odds of being affected when lacking the genetic variant (see **Table 8.6** for a worked example).

TABLE 8.6 A WORKED EXAMPLE OF THE ODDS RATIO IN CASE-CONTROL STUDIES

HLA-Cw6 status	Number of cases (with psoriasis)	Number of unaffected controls	Odds of being affected		Odds ratio
Present	900	330	900/330	→	(900/330) ÷(100/670) =
Absent	100	670	100/670		(900/330) × (670/100) = 18.27

The odds ratio is the odds of being affected when possessing a specific genetic variant divided by the odds of being affected when lacking the genetic variant. In this entirely hypothetical example, we imag-ine a case-control study of psoriasis in which 1000 affected individuals (cases) and 1000 unaffected controls have been typed for the HLA-Cw6 marker, giving the calculation in the final column.

As described in **Box 8.3** and in Section 8.3, certain HLA variants have been identified to be the largest genetic contributors to a wide variety of important autoimmune disorders, and the small HLA region, just over 3 Mb long, has by some distance, the highest density of significant disease associations in our genome.

BOX 8.3 HLA ASSOCIATIONS WITH AUTOIMMUNE DISORDERS

The human major histocompatibility complex (MHC; also called the HLA complex) extends over 3.3 Mb at 6p21.3. It contains many genes that function in the immune system, notably HLA genes. Some of the HLA genes are *classical* HLA genes that make highly polymorphic cell surface proteins involved in cell-mediated immune responses (Box 4.3 on page 105–6 gives HLA nomenclature and a simplified HLA gene map). Classical HLA genes work in cell-mediated immunity to signal the presence of cells infected by viruses (or other intracellular pathogens) to suitably discriminating T cells, thereby initiating an immune response to kill the infected cells.

All proteins within our cells (whether of normal host origin or from intracellular pathogens such as viruses) undergo turnover, whereby the proteins are degraded to peptides within the proteasome. The resulting peptides are bound by newly synthesized HLA proteins and are then transported to the cell surface so that the HLA–peptide complex can be recognized by a specific T-cell receptor on the surface of T cells (**Figure 1A**).

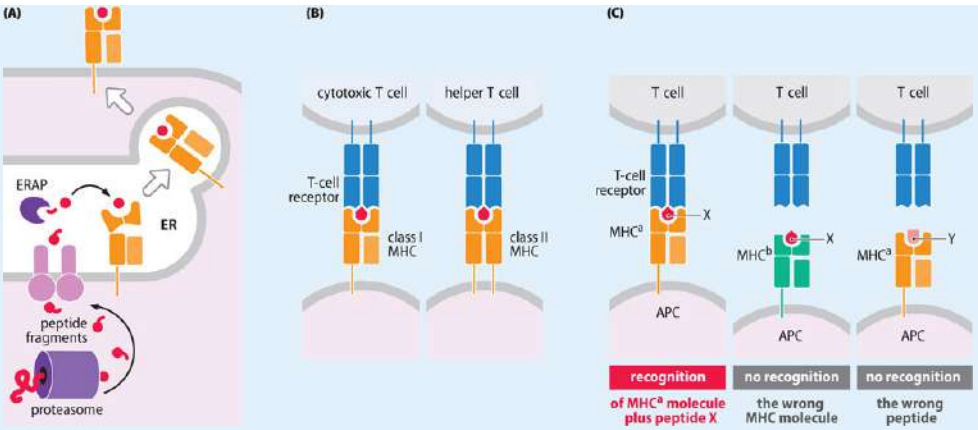


Figure 1 MHC-peptide binding and MHC restriction in antigen presentation. (A) Class I MHC proteins (class I HLA proteins in human cells) serve to bind peptides and display them on the cell surface. The peptides are produced by the degradation of any protein synthesized within the cell (a host cell protein or one made by an intracellular pathogen). Peptide fragments are produced within the proteasome and transported into the endoplasmic reticulum. Here they are snipped by an endoplasmic reticulum aminopeptidase (ERAP) to the proper size needed for loading on to a partly unfolded class I HLA protein. Once the peptide has been bound, the HLA protein completes its folding and is transported to the plasma membrane with the bound peptide displayed on the outside. (B) Receptors on cytotoxic T cells bind class I MHC–peptide complexes; those on helper T cells bind class II MHC–peptide complexes. (C) **MHC restriction**. Each T cell has a cell-specific receptor whose function is dependent on co-recognition of two molecules on the surface of an antigen presenting cell (APC): the combination of a *specific* peptide bound by a *specific* MHC protein. The T cell illustrated here is specific for a particular peptide X bound to a specific MHC allele (an imaginary allele that we designate here as MHC^a as shown on the left). If instead the APC had peptide Y bound to MHC^a, or peptide X bound to MHC^b, it would not be recognized by the same T cell (Adapted from Murphy K [2011] *Janeway's Immunobiology*, 8th ed. Garland Science.)

Immune tolerance ensures that self-peptides (originating from normal host proteins) do not normally trigger an immune response. At an early stage in thymus development, T cells with receptors that recognize self-peptides bound to HLA are eliminated; thereafter T cells are normally focused on nonself ("foreign") peptides (such as those from pathogens). Different T cells in a person contain different T-cell receptors to maximize the chance that a nonself peptide presented by an HLA protein can be recognized. When that happens, T cells are activated to mount an immune response (see **Figure 1B,C**).

Viruses readily mutate in an attempt to avoid triggering immune responses, and the number of potential foreign peptides is huge. This explains why T-cell receptors are genetically programmed, like antibodies, to be extraordinarily diverse (detailed in Section 4.5). HLA proteins vary in their ability to present specific peptides for recognition and so they, too, are selected to be highly polymorphic.

In autoimmune disorders there is a breakdown in the ability to distinguish self from nonself. As a result, cells in the body can be attacked by *autoantibodies* and by

autoreactive T cells that inappropriately recognize certain host antigens (autoantigens). In diseases such as type 1 diabetes, rheumatoid arthritis, and multiple sclerosis, activated T cells kill certain populations of host cells (such as insulin-producing pancreatic beta cells in type 1 diabetes). In autoreactive T-cell responses, host peptides (autoantigens) are presented by HLA proteins that may differ in their ability to bind the autoantigen. As a result, specific HLA antigens are associated with disease.

At the classical HLA loci, large numbers of alleles can be typed (previously, as serological polymorphisms by using panels of antisera; more recently as DNA variants). HLA–disease association studies involve typing HLA gene variants in affected individuals and controls and calculating the frequencies of a specific antigen or DNA variant in the two groups. This allows calculation of the odds of a disease occurring in individuals with or without a particular genetic variant, and calculation of odds ratios (Table 1).

TABLE 1 EXAMPLES OF xsHLA DISEASE ASSOCIATIONS IN THE NORWEGIAN POPULATION

Disorder	Class of HLA antigen	HLA antigen frequency		Odds ratio *
		Affecteds	Controls	
Ankylosing spondylitis	HLA-B27	>0.95	0.09	69.1
Celiac disease	HLA-DQ2and-DQ8	0.95	0.28	15.4
Multiple sclerosis	HLA-DQ6	0.59	0.26	4.1
Narcolepsy	HLA-DQ6	>0.95	0.33	129.8
Psoriasis	HLA-Cw6	0.87	0.33	13.3
Rheumatoid arthritis	HLA-DR4	0.81	0.33	3.8
Type 1 diabetes	HLA-DQ8and-DQ2	0.81	0.23	9.0
	HLA-DQ6	<0.1	0.33	0.22

All of the associations shown here indicate disease risk, except for the negative association of HLA-DQ6 and diabetes (with an odds ratio<1).That is, HLA-DQ6 is a protective allele: carriers have less risk of type 1 diabetes than the general population. (HLA antigen frequency courtesy of ErikThorsby.)

* Table 8.6 shows how odds ratios are calculated.

From Table 1 it is clear that possession of certain HLA antigens confers a substantially increased risk for certain disorders, and the odds ratios can be very impressive. If, for example, you carry an HLA-B27 antigen, you have a much-increased risk of developing ankylosing spondylitis (a form of inflammatory arthritis affecting the joints of the lower back). But HLA-B27 is merely a *susceptibility* factor: although the odds ratio approaches 70, only 1–5 % of individuals with HLA-B27 develop ankylosing spondylitis.

Linkage disequilibrium as the basis of allelic associations

Associations between genetic variants and disease can be caused by different factors, both genetic and nongenetic. In the former case, population substructure and history are important. As previously considered in Section 5.4, human populations within countries, regions, and cities are often stratified into different groups (organized along ethnic, cultural, and religious lines) whose members preferentially mate within the group rather than with members of another group. As a result of population *stratification*, different subgroups within a broad population often have significantly different frequencies of a genetic variant, and this can confound genetic analyses. To minimize problems arising from population stratification, association studies need controls with the same type of population ancestry as those with the trait being studied.

Genetic variants associated with disease might be directly involved in pathogenesis, or they may be tightly linked to a disease-susceptibility allele. In the latter case the haplotype containing the genetic variant and the disease susceptibility allele has a higher frequency than would be predicted from the individual frequencies of the genetic variant and susceptibility allele. This is an example of **linkage disequilibrium**, the nonrandom association of alleles at two or more loci.

As a concept, linkage disequilibrium describes *any* nonrandom association of alleles at different loci; in practice, the alleles are at very closely linked loci. For example, in populations originating in northern Europe linkage disequilibrium is often evident for alleles at closely neighboring HLA genes, such as the *HLA-A* and *HLA-B* loci which are separated by 1.3 Mb of DNA at 6p22. Thus, in the population of Denmark the frequencies of *HLA-A1* and *HLA-B8* are 0.311 and 0.237, respectively, but the frequency of the *HLA-A1–HLA-B8* haplotype is 0.191, more than 2.5 times the expected value of 0.074 (= 0.311 × 0.237).

Linkage disequilibrium might occur if a particular combination of alleles at neighboring loci were positively selected because they worked together to confer some advantage. However, linkage disequilibrium may often simply reflect reduced recombination between loci. This can happen in areas of the genome where there are low recombination rates. We now know, for example, that the HLA complex, the human major histocompatibility complex, is a region of low recombination (with 0.49 cM per Mb of DNA, compared to a genome wide average of 0.92 cM/Mb).

When a new DNA variant emerges by mutation it will show very tight linkage disequilibrium with alleles at very closely linked loci. The linkage disequilibrium will be gradually eroded by recombination, but that will take a very long time for any locus that is physically very close to the locus with the new mutation.

Sharing of ancestral chromosome segments

Association studies depend on linkage disequilibrium, which in turn reflects shared chromosome segments in large numbers of people because of a very distant common

ancestor. Throughout this book we talk about families—groups of people who share large parts of their genomes because of common recent descent. We speak about people being *related* to each other, but we are all related if we go far enough back in history.

What we mean by "related" is having a *known* common ancestor (usually one that can be identified within the previous four generations). And when we say that two persons are unrelated, we generally mean that they do not have any great-grandparent in common, and that they are unaware of any more distant common ancestor. So-called unrelated people do, however, share small common chromosome segments that they have inherited from more distant common ancestors. If the common ancestor lived a long time ago, each shared segment will be quite small but will be shared by a large number of descendants (see **Figure 8.13A** for the principle).

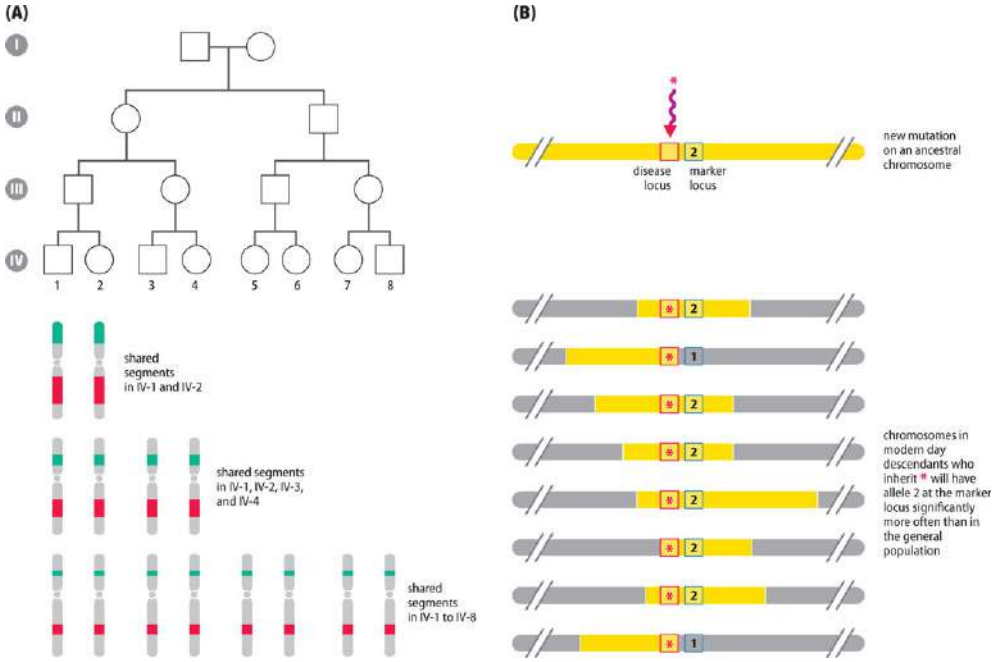


Figure 8.13 Shared ancestral chromosome segments and linkage disequilibrium in the immediate vicinity of an ancestral mutation. (A) The more distant a common ancestor is, the smaller each shared chromosome segment will be, but the larger will be the number of people who share it. In this highly idealized representation, sharing of two chromosome segments on a single chromosome extends to all eight individuals in generation IV (with common great-grandparents). The extent of the shared region is greatest in sibs, but progressively decreases the further the individuals are separated from a common ancestor. (B) Linkage disequilibrium around an ancestral mutation that confers disease susceptibility. Upper panel: imagine a newly emergent mutation (red asterisk) appeared on a chromosome that had a minor (infrequent) allele 2 at a very closely linked SNP locus where allele 1 is the major allele. Lower panel: after passing down through multiple generations, meiotic homologous recombination will ensure that most of the original chromosome (yellow) will have been replaced by segments from other copies of the chromosome (gray). Descendants who inherited the part of the ancestral chromosome with the disease-susceptibility variant have an increased chance of having allele 2 at the very closely linked marker locus. Affected individuals will therefore have a significantly

higher frequency of allele 2 at the marker locus than the general population, or than control unaffected individuals. (Adapted from Ardlie KG et al. [2002] *Nat Rev Genet* 3:299–309; PMID 11967554. With permission from Macmillan Publishers Ltd.)

Shared ancestral chromosome segments can explain linkage disequilibrium. Shared segments contain loci that have not been separated by recombination, and so there is a nonrandom association of alleles at linked loci within such segments. By chance, an ancestral chromosome segment might contain an allele that confers susceptibility to a complex disease. In that case, people living now who suffer from the same disease would tend to share that chromosome segment (Figure 8.13B).

Usually, the susceptibility allele is neither necessary nor sufficient to cause the complex disease; not everyone with the disease will have that allele, and not everyone with the susceptibility allele will have the disease. But, overall, people with the disease are more likely than unaffected people to have that ancestral chromosomal segment. This is the underlying principle that makes disease association studies possible.

Linkage disequilibrium decreases very rapidly with distance between alleles. If genomewide association studies are to be carried out successfully, a marker map with a very high density is therefore needed to cover the genome. On the basis that ~1 nucleotide in 300 is polymorphic, a total maximum of ~10 million single nucleotide polymorphism (SNP) loci are potentially available and the International HapMap Consortium and follow-up studies have mapped and genotyped several millions of SNP loci in different human populations, providing an excellent resource for genomewide association studies. GWA studies have however saved costs by using SNP chips (microarrays) that have just a subset of the total SNP markers (often just 500 000 or a million SNP markers). As we explain later, when we describe how GWA studies are conducted, it has been possible to use statistical methods to infer genotypes at untested SNP loci (ones not represented on the SNP chips used in GWA studies).

The HapMap data show that our nuclear genome is a mosaic of small blocks of sequence, **haplotype blocks**, in each of which there is very limited genetic diversity. The blocks vary in size, but average ~5 kb, and at most chromosomal locations most genomes have just one out of only 3–5 different ancestral blocks. That does not, of course, mean that most of us are descended from 3–5 remote ancestors (at a neighboring block most genomes may again have one of only 3–5 ancestral blocks but they would be inherited from a different 3–5 remote ancestors, and so on; our remote ancestry is with *populations*, not individuals). **Box 8.4** provides more details and visual representations of haplotype blocks.

BOX 8.4 HAPLOTYPE BLOCKS AND THE INTERNATIONAL HAPMAP PROJECT

Initial attempts to define ancestral chromosome segments began with high-resolution mapping of haplotype structure in defined small genome regions in populations of European ancestry. The results suggested that our nuclear DNA might be composed of defined blocks of limited haplotype diversity (**haplo-type blocks**). Figure 1A illustrates an example—a haplotype block 84 kb long that spans most of the *RAD50* gene at 5q31. Eight common SNP loci were genotyped in this block, and two alleles at each of eight SNP loci means the potential for $2^8 = 256$ different haplotypes. Yet, within this block, almost every chromosome 5 that is tested has 1 of only 2 out of the 256 possible haplo-types—either the orange haplotype in Figure 1A (which we can represent by listing the nucleotides at the eight consecutive SNP loci as GGACAACC) or the green haplotype (AATTCGTG).

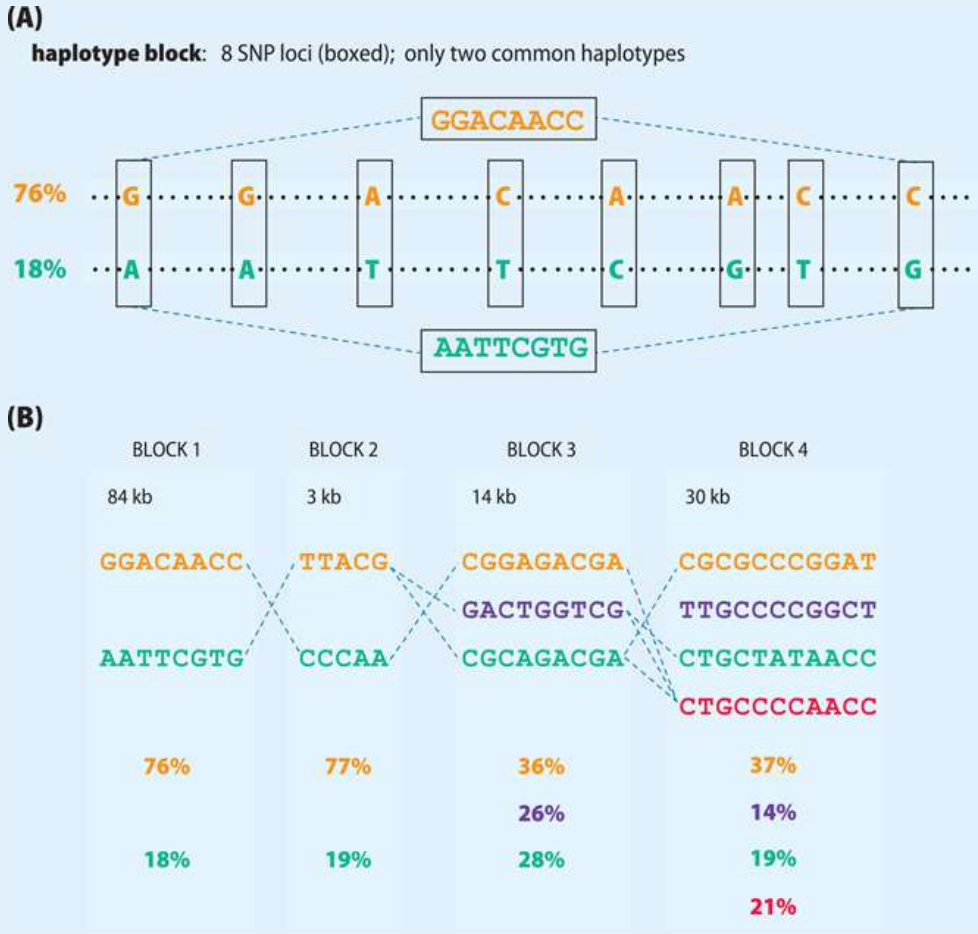


Figure 1 **Haplotype blocks.** (A) Genotyping at eight SNP loci (vertical boxes) spanning most of the *RAD50* gene at 5q31 reveals an 84 kb haplotype block. Just two haplotypes account for the vast majority (76 % and 18 %, respectively) of the chromosomes 5 from a sampled European population. (B) Adjacent haplotype blocks at 5q31. The 84 kb the block from panel (A) above, is represented as block 1 here. Neighboring blocks 2, 3, and 4 were genotyped at respectively five, nine, and eleven SNP loci and had between two and four haplotypes shown in different colors at population frequencies given at the bottom. Dashed blue lines signify locations where >2 % of all chromosomes 5 are seen to switch from one common haplotype to

another. (Adapted from Daly MJ et al. [2001] *Nat Genet* 29:229–232 PMID 11586305. With permission from Macmillan Publishers Ltd.)

The low haplotype diversity is apparent in adjacent haplotype blocks. In Figure 1B, block 1 (the same block that is shown in Figure 1A) and the neighboring block 2 are dominated by two haplotypes, and the next two blocks by three and four haplotypes, respectively. It suggests that the DNA in block 1 was contributed mostly by two ancestors, and that in blocks 2, 3, and 4 by a different set of two, three, and four ancestors, respectively.

The International HapMap project set out to make comprehensive maps of linkage disequilibrium in the human genome. The project began by genotyping common single nucleotide polymorphisms (SNPs) in samples from four populations: the Yoruba from Nigeria (YRI); a white population from Utah, USA, descended from northern and western Europe (CEU); Han Chinese from Beijing (CHB); and Japanese from Tokyo (JPT). Hapolotype maps were constructed by genotyping 3.1 million SNPs (or about one every kilobase).

The HapMap project confirmed that humans show rather limited genetic variation (by comparison, chimpanzees show very much more genetic diversity). At a fairly recent stage in population history, the human population was reduced to a very small number—perhaps just 10 000 or so individuals—that remained quite constant until comparatively recently. First agriculture, and then urbanization led to a very rapid massive expansion in population size to the current eight billion individuals. As a result, about 90 % of the genetic variation in humans is found in all human populations.

Overall, about 85 % of our nuclear genome is a mosaic structure, composed of haplotype blocks. The average size of the haplotype blocks in the populations of European and Asiatic ancestry was 5.9 kb with an average of about 3.6 different haplotypes per block. In the Yoruban population there were an average of 5.1 different haplotypes per haplo-type block and the blocks averaged 4.8 kb in size (all human populations originated in Africa, and African populations have greater genetic diversity). Note, however, that the number, size, and identity of blocks depends on the statistical criteria used to define a block.

How genomewide association studies are carried out

Genomewide association (GWA) studies (or **GWAS**) began to really take off in the mid-2000s because of two technological developments. First, the International HapMap Project delivered hundreds of thousands and then millions of mapped SNP (single nucleotide polymorphism) loci. Secondly, by the mid-2000s the extension of microarray technology allowed the automated geno-typing of huge numbers of SNPs across the genome. We described the principles of microarray technology previously in Figure 3.9 at the end of Section 3.3. In the case of whole genome SNP microarrays, the microarrays carry

oligonucleotides specific for each allele at many hundreds of thousands of SNP loci across the genome, plus controls.

GWA projects were designed to identify *common* variants (it was assumed that common complex diseases are predominantly caused by common variants). The bulk of GWA studies have therefore focused on case-control studies in which panels of affected individuals and matched controls are genotyped at hundreds of thousands of common SNPs (where the minor allele usually has a frequency of at least 0.05). SNPs are then identified in which allele frequencies are significantly different in cases than in controls (Figure 8.14).

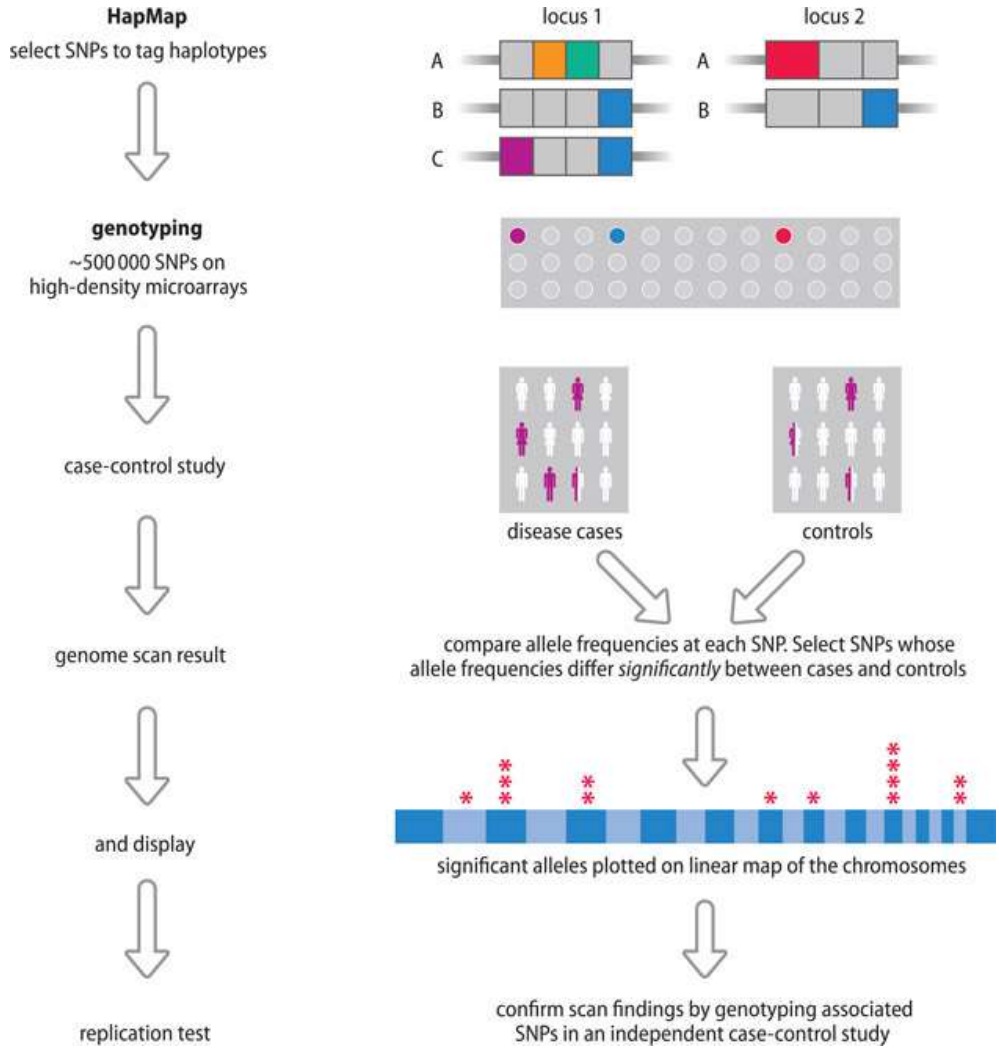


Figure 8.14 Carrying out a genomewide association scan using SNP chips (microarrays). Using HapMap data (which map linkage disequilibrium across the human genome), representative single nucleotide polymorphisms (SNPs) are selected that will differentiate (or *tag*) the common haplotypes at each locus. In this example, three common haplotypes (A, B, and C) at locus 1 are tagged by four SNPs (with strong color if present, or gray if absent). But just two SNPs (purple and blue) are sufficient to discriminate between the three haplotypes. Similarly, the two haplotypes at locus 2 can be distinguished by either the red (chosen here) or the blue SNP The *tag SNPs* are then genotyped in disease cases and controls using microarrays, and the allele

frequencies for each SNP are compared in the two groups. SNPs associated with disease at an appropriate statistical threshold are genotyped in a second independent sample of cases and controls to establish which of the associations from the primary scan are robust. (Adapted from Mathew CG [2008] *Nat Rev Genet* 9:9–14; PMID 17968351. With permission from Macmillan Publishers Ltd.)

Statistical thresholds and data visualization

SNP microarray hybridization typically involves many hundreds of thousands of parallel DNA hybridizations, one for each of the fixed oligonucleotides. Because such huge numbers of hybridization tests are being carried out, stringent statistical significance thresholds are required to assess the significance of individual hybridization results. In order to set a more stringent genome-wide significance threshold, the standard P value of 0.05 is divided by the number of tests carried out. If the microarray hybridization involves one million different hybridization assays, for example, a stringent P value would then be $0.05/1\ 000\ 000 = 5 \times 10^{-8}$. One consequence of having such a stringent cut-off is that true weak positives might not be recorded. We consider the significance of that below.

The genotype test statistics are calculated for each variant and referenced against statistics expected under the null hypothesis of no disease association. The data can be visualized in different types of plot, notably Manhattan plots as shown in **Figure 8.15**.

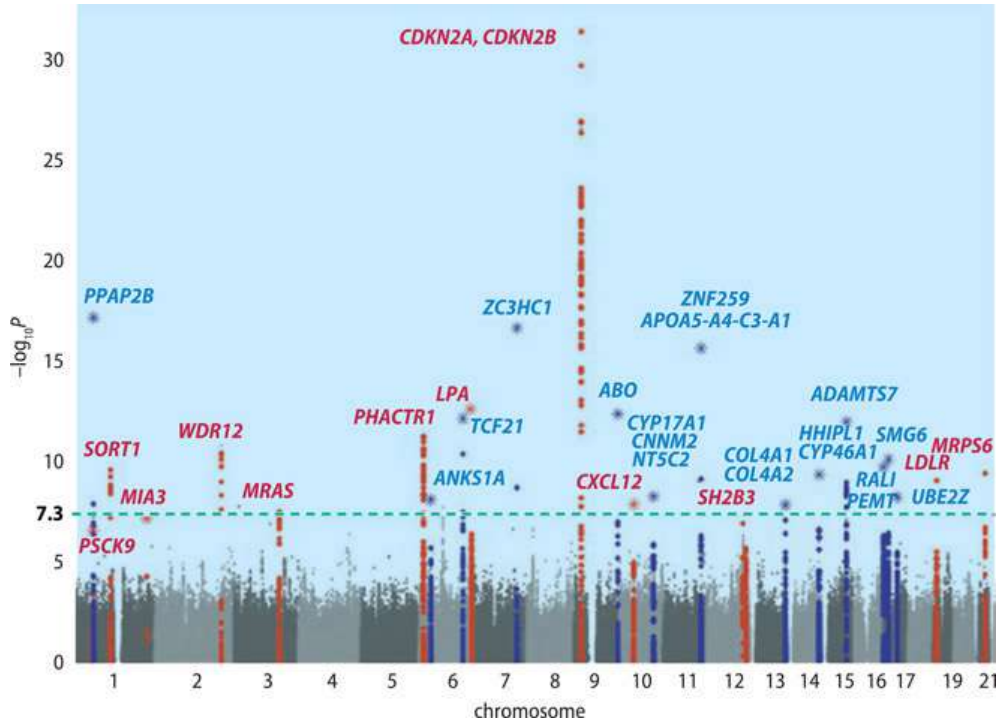


Figure 8.15. Visualizing genomewide association (GWA) data. The most common representation is a genome-wide Manhattan plot (think of skyscrapers). It displays GWA signals according to their genomic

positions (horizontal scale) and statistical significance (vertical scale; the negative $\log_{10}P$ scale helps reveal signals of particular interest). This plot is from a large study of coronary artery disease; newly discovered disease-susceptibility loci are in blue, and previously discovered ones in red. The dashed green horizontal line at position 7.30 on the vertical scale indicates the threshold of statistical significance (corresponding to $P = 5 \times 10^{-8}$ in this case). The most significant associations were with previously recorded SNPs in the immediate vicinity of the closely neighboring *CDKN2A* and *CDKN2B* genes at 9p21. (From Schunkert H et al. [2011] *Nat Genet* 43:333–338; PMID 21378990. With permission from Macmillan Publishers Ltd.)

The need for phasing and imputation

The output from standard SNP microarray GWA studies is in the form of geno-types. However, as explained above, population associations are explained by shared ancestral chromosome segments. Accordingly, disease risk is thought to be defined by *haplotypes*, not genotypes. In modern GWA studies the procedure known as **phasing** is employed to arrange genotypes of neighboring loci into haplotypes. Interested readers can find more detail in the Browning and Browning (2011) reference under Further Reading. There are two immediate benefits of phasing. First, it provides haplotype-disease associations, which can be expected to be more accurate than the simple allele-disease associations obtained when analysing the raw genotype data. Secondly, phasing is essential for being able to carry out an important process widely used in GWA studies: genotype imputation.

In statistics, **imputation** is any process designed simply to estimate missing data. SNP chips used in GWA studies usually allow genotyping of hundreds of thousands or sometimes a million SNP loci. That is a rather small fraction of the SNP loci in the human genome (around one nucleotide is 300 is polymorphic, giving ~10 million or so SNP loci across the genome). That is, there is a lot of missing data that could potentially be mined.

Genotype imputation seeks to estimate the identity of the alleles at many of the untested SNP loci, by using reference genotype data and by taking advantage of linkage disequilibrium. Extensive information is available from reference human haplotypes across the genome obtained by various projects where extensive genotyping has been carried out. The initial reference human haplo-types came initially from the HapMap project, then subsequently from other sources: the 1000 Genomes, UK10K or TopMed projects, and then the Haplotype Reference Consortium (HRC), the most widely used imputation reference panel.

By looking for ancestrally shared regions of chromosome between a GWAS sample and individuals in the reference panel, alleles can be inferred with a very high degree of probability because haplotype sharing can extend over signifi-cant regions. When a typical sample of European ancestry is compared to haplo-types in the reference panels, for example, shared stretches of >100 kb in length are often identified. It is not easy to represent this visually, but **Figure 8.16** gives the general idea.

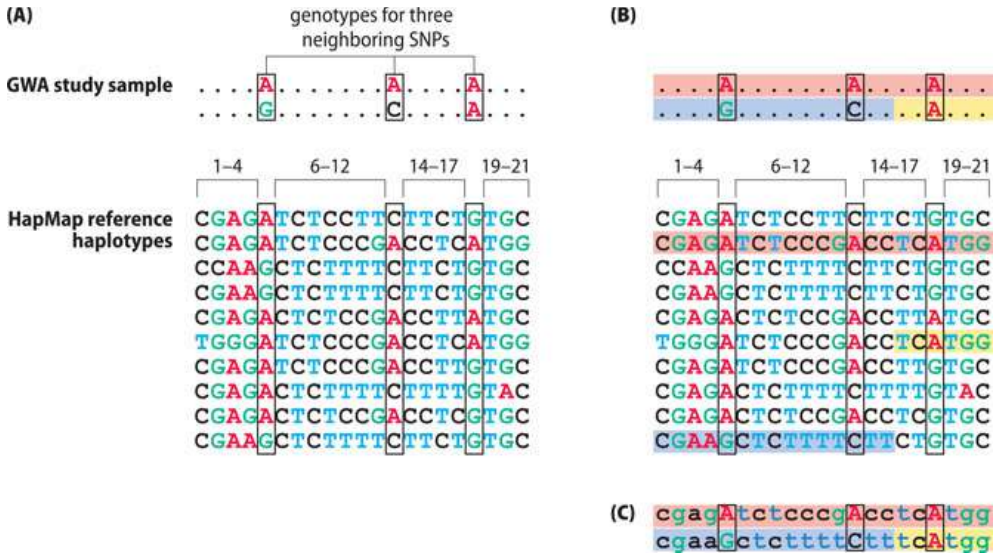


Figure 8.16 Genotype imputation in GWA studies. (A) The observed data. Genotypes obtained in the imagined GWA study here are given for three SNP loci that happen to be represented in the SNP microarray out of a total of 21 SNP loci in the immediate region (the dots represent the missing data, that is, the unknown alleles at the untested 18 SNP loci). For the same region various haplotypes are available from the HapMap where alleles at all 21 SNP loci have been identified. (B) Identifying regions of chromosome shared between a study sample and individuals in the reference panel (sharing is indicated by color shading). (C) The genotypes of the original three SNP loci (shown in upper case letters in the black open boxes), have been supplemented by the haplotype sharing information from (B) so as to reconstruct haplotypes for all 21 loci, even although only 3 of the 21 SNP loci were tested in the GWA study.

Genotype imputation is helpful in finding new disease-associated loci (simply by imputing the additional SNP loci, you get extra statistical power for the study). It can also be useful in fine mapping. A signal indicating association may be obtained with one of the tested SNP markers. But there will be other untested SNP markers nearby. If additional markers in this region are imputed a marker may be found that is more highly correlated with the disease; it would then become a priority for being included in replica studies. But by far the most important contribution made by genotype imputing to GWA studies is to enable large *meta-analyses*, as described below.

Dealing with confounding sample structure

Elements of sample structure, such as population subgroups, family relatedness and/or cryptic relatedness, are important confounders in GWA studies. (A *confounder* is a statistical term for a variable factor that influences both an independent and a dependent variable, such as, respectively genotype and disease incidence.) For example, when there are subgroups in the population with ancestry differences, one subgroup of the population may accidentally be

overrepresented in the study sample and another subgroup may be underrepresented, causing spurious (false-positive) disease associations.

Family relatedness can also confound studies. Ideally, individuals in a population under study would not be (recently) related, but that assumption may often not be true. Various *linear mixed models* perform well in dealing with confounding sample structure in GWA studies. Interested readers can find detailed explanations of the approaches used in recent reviews such as at PMID 24473328 and PMID 25033443.

The importance of very large well-designed studies and meta-analyses

Initial promising GWA subchromosomal locations need to be confirmed. To do this, candidate SNPs of high statistical significance are genotyped in an independent replication panel. In addition to low *P* values, extra confidence is obtained when the same location is replicated in independent studies on different populations.

Initial GWA studies were of comparatively small scale and more suited to mapping the comparatively few susceptibility factors of quite large effect. To map the more numerous susceptibility factors with more modest effect, large numbers of cases were needed. The template for successful large GWA studies was established by the Wellcome Trust Case Control Consortium (WTCCC), a consortium of British researchers who maximized their chances by pooling their individual collections of cases and controls in well-designed studies. In a landmark paper in 2007, the WTCCC reported considerable success in mapping susceptibility to seven complex diseases with 2000 cases for each disease and a common set of 3000 controls. There followed an explosion in the number of similar GWA studies and very considerable success, as many initial findings have been replicated and confirmed.

Human SNP chips made by different companies often show little overlap in the SNP markers chosen to be represented on the chips. Direct pooling of the data from research studies that use different types of SNP chip is, therefore, extremely problematic. But this is where imputation is so valuable: by imputing the same sets of alleles from many HapMap reference SNP loci, a common set of genotyping data become available for research studies allowing legitimate larger-scale pooling of the data. Such high-powered **meta-analyses** have been the foundation of the important successes of GWA studies in recent times.

Moving from candidate subchromosomal region to identify causal genetic variants in complex disease can be challenging

An SNP that shows a significant disease association is expected to be closely linked to the "causal variant", the genetic variant that plays a role in causing the disease (by altering how a gene is expressed). However, moving from an associated SNP to a genetically linked causal variant may be extremely difficult for a variety of reasons, as listed below.

- GWAS variants often fall within regions of high linkage disequilibrium. As a result, many non-causal variants are inherited together with the causal variant. Because they, too, are on the same shared small chromosome segment, all of them will also be associated with the disease. Picking the causal variant normally means having to sift through multiple variants, and the candidate region may sometimes be quite large (when the local region shows relatively low recombination).
- Unlike in monogenic disorders, there is the problem that the causal variant is often not causative! It will be absent in a proportion of people with the disease and will be found in many normal people—the "causal" variant is merely a susceptibility factor. Take, for example, the hypothetical association of allele *A*1* with disease D. Imagine that *A*1* has a frequency of 0.457 in 2000 affected cases and a frequency of 0.361 in 2000 controls. That might seem a small difference, but because of the large number of samples genotyped it would be highly significant.
- Most GWAS variants, and even variants in strong linkage disequilibrium with the lead variant, are located in noncoding DNA regions, making it difficult to infer their importance. A noncoding variant might be expected to affect some noncoding regulatory DNA sequence (which might be poorly understood or previously uninvestigated), causing a change in expression of a nearby gene that might be limited to specific tissues or cell types. Occasionally, causative variants in coding DNA are found, but like the variants affecting regulatory elements, they might be expected to have a mild effect (for a causal variant to be common in the population it must be a comparatively ancient mutation: it cannot have been exposed to strong purifying (negative) selection otherwise it would have been eliminated by now). A causal variant in coding DNA is also likely to be subtle (such as a mild missense mutation that might not be readily recognized). There are exceptions: easily identified frameshifting and nonsense mutations may sometimes be implicated that would be expected to result in retention of most of the normal protein sequence (for an example, see the Crohn's disease-associated *NOD2* 3020insC variant described in the legend to Figure 8.11A).

Towards identifying causal variants and disease-susceptibility loci

As mentioned above, a disease-associated SNP will have assorted neighboring variants mapping to the same haplotype block; they, too, will be disease-associated. But because the borders of haplotype blocks are not precisely defined (the linkage disequilibrium is always significantly <100 %), there may be differences in the degree of association of the different variants within a block. Within each block are a limited number of haplotypes; variants that happen to be present in multiple haplotypes will be much more strongly associated than those on a single haplotype. Ideally, what we would hope to find is a peak area where a few

variants show especially strong association, against a general background of association for the variants within a critical region (which can encompass adjacent haplotype blocks). Readers who may be interested in the statistical fine-scale mapping methods to home in on causal variants can find a relevant review at PMID 29844615. As the search draws closer to identifying the causal variant and the locus of the disease susceptibility factor, additional approaches can be used.

- *Candidate genes*. If a previously studied gene in the immediate neighbour-hood appears to be related to previously identified disease-susceptibility loci at other genomic locations, or is involved in similar biological pathways, it immediately becomes a candidate susceptibility factor locus. It would then be a priority for intensive studies, as listed below.
- *Expression studies.* In principle, genes in the immediate vicinity, or a candidate gene if identified, can be studied to determine if their expression patterns appear to be significantly affected when comparing the presence or absence of the causal variant in laborious wet-lab experiments.
- *Sequencing studies.* Possible coding DNA variants can be screened with programs that assess the likely effect of amino acid substitution on the structure and function of the predicted protein, using programs such as PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (https://sift.bii.a-star.edu.sg/), and PROVEAN (http://provean.jcvi.org/index.php).

The limitations of GWA studies and the issue of missing heritability

Early linkage studies identified subchromosomal locations for some important genetic variants underlying complex disease, but the overall returns from later studies were minimal. Genomewide association (GWA) studies have had much greater success, and meta-analyses, in which the data from individual large disease association studies are pooled, provide even more power to detect significant associations, and have been especially valuable. Thousands of significant disease associations have been obtained, very many of which have been replicated.

Despite initial high hopes, the common disease variants identified by GWA studies have generally been of very weak effect—often with an odds ratio of 1.2 or less. Exceptions include some novel factors strongly predisposing to age-related macular degeneration (a leading cause of vision loss in older adults due to progressive deterioration of a central region of the retina). But many of the variants with high odds ratios were identified in the pre-GWAS era (such as *APP*, the amyloid precursor protein gene, in Alzheimer disease, the common *NOD2* alleles in Crohn's disease, and especially HLA alleles that remain, by some distance, the strongest known genetic variants in autoimmune disorders).

The “missing heritability” problem

Once the effect size of an individual GWAS variant is known, its contribution to the heritability can be calculated. By doing this type of calculation for each GWAS variant associated with a specific disease, an aggregate estimate can be obtained for the combined contributions of all of the GWAS variants. This bottom-up approach, however, has frequently given much lower heritability estimates than those from top-down family studies (such as the twin studies described above). That then posed some important questions. Are the GWA studies missing something? And how can we account for the “missing heritability”?

Many different explanations have been put forward to explain the missing heritability problem. However, current thinking identifies two major reasons, resulting from self-imposed constraints in the statistical analysis of GWA studies and in the experimental GWA study design—see below.

- *The statistical significance cut-off for association*. The statistical cutoffs were chosen on the assumption that most association signals represent noise, and that only a few can be real (this assumption was based on what used to be a widespread oligogenic view of complex disease; as explained below, truly polygenic contributions may be the reality for individual complex diseases).
- *The filtering out of low frequency alleles*. The experimental design of GWA studies also requires a cut-off in terms of permitted allele frequency: only common SNP alleles (segregating with frequencies >5 % in the population) have been eligible as standard GWAS alleles. Rarer alleles, which collectively might have made an important contribution to the phenotype variation, were excluded for technical reasons (because of the reduced probability of linkage with causal variants and low statistical power).

The limitations of GWA studies

The motivation for GWA studies was, largely, for two reasons: to improve the prediction of disease risk (identifying individuals as being at higher risk for a specific disease should hopefully enable preventative measures and/or better targeting of clinical resources); and to provide a systematic approach to identify genes involved in pathogenesis, offering more possibilities for developing drugs and more effective therapies.

Early GWA studies were underpowered and were largely limited to identifying common variants of quite large effect, but thereafter increased sizes of individual GWA studies and then combinatorial meta-analyses were able to extend the scope of GWA studies very significantly. GWA studies are now widely viewed to have been technically very successful,

delivering many thousands of unique, robust associations between common variants and common diseases. Nevertheless, they have been widely believed to have underperformed for a variety of reasons—see **Table 8.7**.

TABLE 8.7 FIVE REASONS WHY GWA STUDIES HAVE NOT BEEN AS SUCCESSFUL AS HOPED

Reason	Explanation
Common variants almost always have very small effects	The genetic contribution to disease risk is often made up of an extremely large number of variants with very small effects, often with very low odds ratios (the great majority would need astronomically large study sizes to be definitely implicated).
Missing heritability	The common SNP variants assayed in GWA studies can explain only a part of the heritability. Rare variants and copy number variants also make important contributions (see text).
Limits on resolution	GWA studies are made cheaper by testing a limited percentage of SNP markers and relying on *imputation* to infer, from linkage disequilibrium, the genotypes at most SNP loci. The cost savings are made at the expense of resolution (GWA studies rely on implicating SNP-causal variant haplotypes rather than genotyping all SNP variants).
Causal variants are very difficult to identify	The great majority of common causal variants are in comparatively unstudied noncoding DNA. Time-consuming experimental approaches and/or very large-scale targeted sequencing may be needed to implicate variants and so far, few causal variants have been identified.
Interconnected regulatory networks	Many genuine disease associations may be due to genes that only subtly impact genes in core pathways contributing to pathogenesis.

The value of carrying out future GWA studies of ever greater scale to hunt down variants with very small effects is questionable. Instead, the priority must be to investigate further the thousands of disease associations already obtained. As described above, the work involved in moving from a single disease-associated variant to identify the linked causal variant unambiguously can still be significant.

As described in Section 8.3 GWA studies have been very important in helping us understand the underlying pathology of many common genetic diseases. As yet, however, at the beginning of the 2020s they have not yet had much effect on public health, and have had little clinical utility. GWA-derived polygenic risk scores seemed to have some promise, and we explain the background to them in the final subsection of Section 8.2.

Alternative genome-wide studies and the role of rare variants and copy number variants in complex disease

Standard SNP chips used in GWA studies are not suited for identifying rare variants or copy number variants (which can be important in some disorders). Genomewide DNA sequencing can be readily used to identify rare variants that have one or a small number of nucleotides changed, but is not so suited to identifying many CNV loci (Note: the term *copy number variant* is poorly defined, but in the context of genomewide studies it is commonly used to mean copy number changes in sequences of intermediate length—from 0.1kb to a few megabases—where variation is due to simple deletions or duplications rather than replication slippage). For such large CNVs, microarray-based methods are used such as the CytoScan HD platform (which has 1.9 million copy number markers; interested readers can find a review at PMID 30223503). As described below, rare variants and CNVs have been found to be important in some complex diseases.

The importance of rare variants in complex disease

The proportion of phenotypic variance captured by all common SNPs, the SNP-based heritability, is substantially less than estimates of pedigree heritability. Even when using SNP genotypes imputed from a fully sequenced reference panel to gain additional additive variance, a significant gap remains between SNP-based and pedigree-based heritability estimates. The hypothesis that causal variants are rare, and consequently not well tagged (or imputed) by common SNPs, has recently been tested.

After carrying out whole genome sequencing (WGS) to estimate heritability of height on a large sample of individuals from the Trans-Omics for Precision Medicine (TOPMed) program, Pierrick Wachstein, Peter Visscher and colleagues estimated the WGS heritability for height to be significantly greater than SNP-based heritability estimates but approaching pedigree heritability estimates. The implications of the Wachstein et al. (2022) paper, listed in Further Reading, is that rare variants, particularly those in regions of low linkage disequilibrium, are a major source of the missing heritability in complex traits and disease.

More support for the importance of rare variants in complex disease has come from recent disease studies, such as the study of schizophrenia described by Singh et al. (2022), as listed under Further Reading. After whole exome sequencing of a large panel of schizophrenia cases and controls, ultra-rare coding variants were implicated in ten genes, and genes prioritized from studies investigating common variants were found to be enriched in rare variant risk.

Copy number variants associated with complex diseases

The poor initial returns from early (underpowered) GWA studies prompted alternative studies of copy number variants. Some CNVs are quite common (with a frequency of more

than 1 %) and have been described as *copy number polymor- phisms* (*CNPs*). Although the overall contribution of copy number variation to complex disease susceptibility may not be so very high. CNPs have been found to be associated with a variety of different disorders, often by changing the copy number of genes in a clustered multigene family (see **Table 8.8** for examples).

TABLE 8.8 EXAMPLES OF COPY NUMBER POLYMORPHISMS (CNPs) ASSOCIATED WITH COMPLEX DISEASES

CNPallele	Disease
Deletion upstream of the *IRGM* gene (involved in the innate immune response)	Crohn's disease
Low copy number allele for an intragenic CNV within the *LPA* gene encoding lipoprotein Lp(a); shown in Figure 2.13B.	coronary artery disease
High copy number allele that spans the β-defensin multigene family and so provides extra β-defensin genes (which make antimicrobial peptides that provide resistance to microbial colonization of epithelial cells)	psoriasis
Single copy of the complement *C4* gene (instead of the normal two complement *C4* gene copies)	systemic lupus erythematosus
Low-copy-number *FCGR3A* alleles (notably deletions).The *FCGR3A* gene makes the Fc portion of immunoglobulin G and is involved in removing antigen-antibody complexes from the circulation and in other antibody-dependent responses	

Data from Girirajan S et al. (2011) *Annu Rev Genet* 45:203-226; PMID 21854229.

Rarer copy number variants have been particularly implicated in neuropsychiatric disease, such as in autism and schizophrenia. High-resolution karyo-typing has shown that about 5 % of individuals with autism spectrum disorder have cytogenetically visible chromosome rearrangements. Global screens also suggest that there is a greater load of subcytogenetic common CNPs and rare CNVs in individuals with autism than in controls. Duplications, as well as deletions, contribute to disease. Many of the CNVs are found to occur *de novo;* others are transmitted, sometimes from an unaffected parent.

Inherited CNVs are scarcely more frequent in autism spectrum disorder than in controls, but *de novo* CNVs in affected individuals are generally larger than in controls, and typically about three to six times more frequent. Many CNVs associated with autism spectrum disorder contain multiple genes (**Table 8.9**), even if the pathogenesis might be due to altered expression of a single gene in many cases. Schizophrenia-associated CNVs cover some of the same regions found to be associated with autism, such as 1q21.2 (deletion), 22q11.2 (deletion), and 16p11.2 (duplication), plus large deletions of variable size at the *NRXN1* locus at 2p16.3.

TABLE 8.9 EXAMPLES OF LOCI THAT FREQUENTLY UNDERGO COPY NUMBER VARIATION IN AUTISM SPECTRUM DISORDER (ASD)

Locus size	Genes	Location	Pathogenic allele	Frequency in ASD (%)
0.7 Mb	30 genes	16.11.2	deletion and duplication	0.8
~1Mb	*(PTCHD1* and *PTCHD1AS)*	Xp22.1	deletion; mostly affecting upstream *PTCHD1AS* antisense noncoding RNA	0.5
Variable	*NRXN1*	2p16.3	mostly deletion	0.4
1.4 Mb*	22 genes	7q 11.2	duplication	0.2
2.5 Mb	56 genes	22q11.2	deletion and duplication	0.2
1.5 Mb	14 genes	1q21.1	duplication	0.2
Variable	*SHANK2*	11q13.3	deletion	0.1

* The same region is deleted in Williams-Beuren syndrome. (Data from Devlin B & Scherer SW [2012] *Curr Opin Genet Dev* 22:229-237; PMID 22463983.)

The *de novo* CNVs cannot contribute to heritability, and although the CNVs in autism spectrum disorder and schizophrenia are quite often of large effect, they account for only a small proportion of the observed genetic variance.

The assessment and prediction of risk for common genetic diseases and the development of polygenic risk scores

GWA studies have disappointed by offering poor predictive capacity. Take type 2 diabetes, for example. A study published in 2011 by de Miguel-Yanes et al. (PMID 20889853) reported that the predictive power of genetic data added very little to the predictive power of clinical data. The predictive power was measured using a standard statistical measure, the AUC (the receiver operator **a**rea **u**nder the **c**urve statistic—see **Box 8.5**). The clinical indicators measured were: age, sex, family history, body mass index, blood pressure, blood glucose, HDL cholesterol, and triglycerides, and the AUC from the combined clinical indicators was 0.903. When genotypes from 40 known genetic susceptibility factors were added to the clinical indicators, the combined AUC statistic increased to just 0.906. Many previous studies had reported similar results.

BOX 8.5 ASSESSMENT AND PREDICTION OF DISEASE RISK

We describe genetic testing in detail in Chapter 11. For now, note that two important parameters of any genetic test are its sensitivity (the proportion of all people who have the condition who are identified by the test) and its specificity (the proportion of all people

who do not have the condition in whom the test result correctly predicts absence of the condition).

Identified genetic variants for complex disease susceptibility generally show rather low odds ratios. If genetic testing is ever to have high predictive accuracy in complex disease, a battery of tests would be needed. To measure the prediction accuracy of such testing, receiver-operating characteristic (ROC) curves are used. Here, the test sensitivity is plotted against 1 – specificity (the value of the specificity subtracted from 1.0). The area under the curve (AUC) is a measure of how well the test can distinguish between the tested people who have the condition and those who do not. AUC values range from 0.5 (providing no discrimination between those with the condition and those without it) to 1.0 (perfect discrimination). As shown in **Figure 1**, simulations show that AUC values can increase as more genetic susceptibility factors are included.

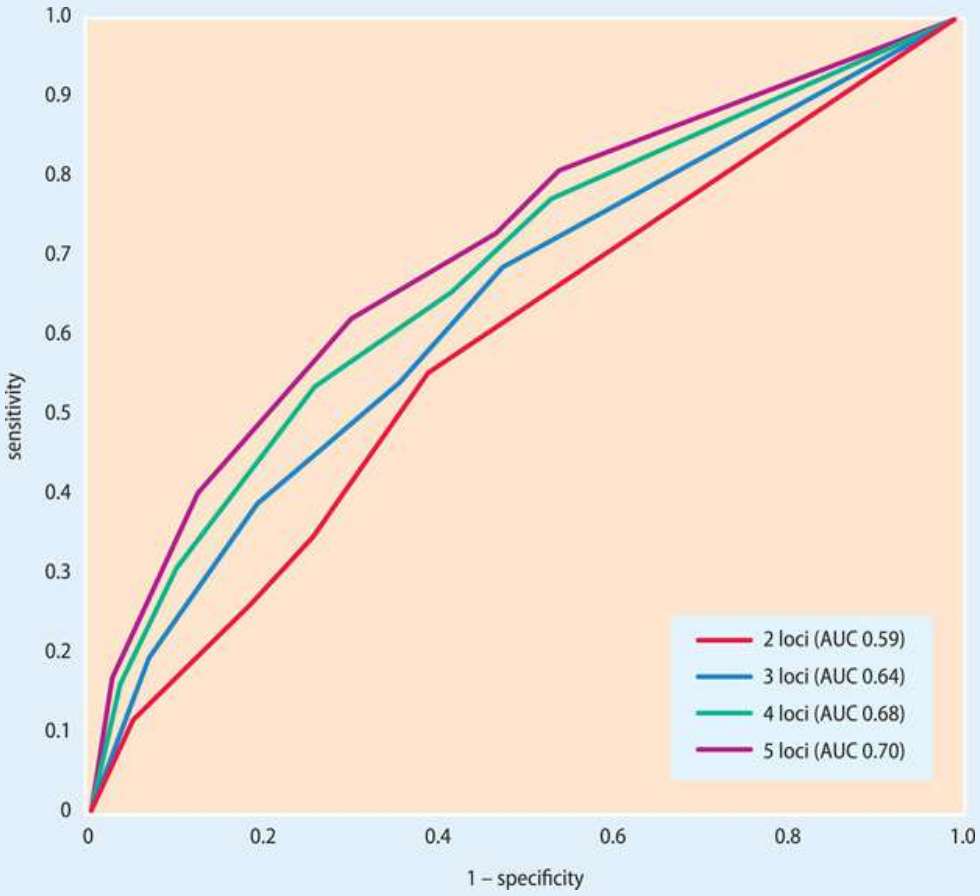


Figure 1 Predictive accuracy of testing for multiple genetic susceptibility factors in complex disease. A receiver-operating characteristic (ROC) curve plots sensitivity against (1 – specificity) for a test. The figure shows ROC curve simulations for testing with two, three, four, or five independent disease susceptibility factors; in this case susceptibility factors 1 to 5 are imagined to have relative disease risks of, respectively, 1.5, 2.0, 2.5, 3.0, and 3.5. (A relative risk of 2.0, for example, would mean that a person with the susceptibility factor has twice the risk of developing the disease compared with a person without it.) Testing

for multiple susceptibility factors can lead to an increase in the area under the curve (AUC); the greater the AUC value, the more discriminating the test. (From Janssens AC et al. [2004] *Am J Hum Genet* 74:585–588; PMID 14973786. With permission from Elsevier.)

Note that a very high AUC predictor may be of little practical use when the disease is quite rare. HLA-B27 is very strongly associated with ankylosing spondylitis, a rare type of chronic arthritis that affects parts of the spine. Despite the very impressive odds ratio of close to 70, and a test sensitivity and specificity each of 99 %, the disease risk conferred by typing positive for HLA-B27 is low (in different populations only about 1–5 % of individuals with HLA-B27 will develop the disease).

For many common complex diseases, even multiple variants identified by GWA studies fail to endow the genetic tests with any great predictive value (most SNP variants have odds ratios of less than 1.3). Current prediction of individual disease risk is not accurate because, for most diseases, only a small proportion of genetic variation in risk between people can be explained by known genetic variants. Type 1 diabetes is at the upper end of the scale: about 70 % or more of familial (genetic) risk can be accounted for by a combination of the major histocompatibility complex (the dominant contributor) and more than 50 additional GWA risk loci. The predictive model has an AUC of close to 0.9, but that is still some distance from what would be desired.

Even if we were to know—and be able to test for—every single genetic risk factor for a disease, the resulting whole genome genetic test would have only partial predictive success, because complex disease is caused by a combination of genetic and environmental factors. Depending on the heritability of a complex disease, the accuracy of genome-wide genetic prediction would have an upper limit of 60–90 % (assuming that we could identify every single genetic variant that affects risk and were able to estimate their effects without error). To obtain truly accurate testing in complex disease, environmental factors need to be taken into account.

Polygenic risk scores

The paradigm of establishing risk of a complex genetic disorder by testing individuals at loci known to affect risk was first challenged in 2009 by David Evans and Shaun Purcell and their collaborators. The question asked was: why confine ourselves to testing *known* risk factors when we have the technology from GWA studies to genotype individuals at huge numbers of loci across the genome. So began the new concept of using GWA-derived *genome-wide* genotypes to construct what was initially called polygenic scores (PGS) but came to be called **polygenic risk scores** (PRS). A typical procedure comprises the three steps shown in **Figure 8.17A**.

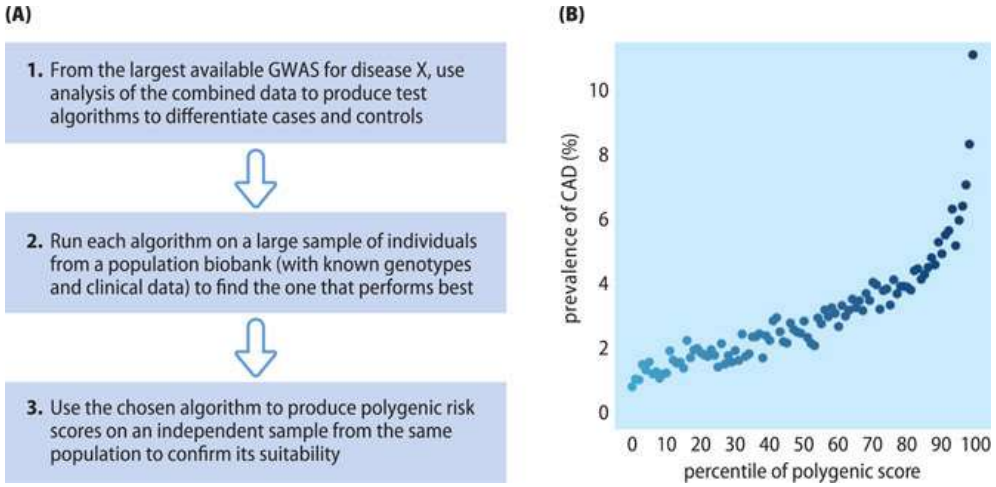


Figure 8.17 Polygenic risk scores: study design and an example. (A) A typical protocol. In step 1 the "combined data" means data on both cases and controls. In step 2, the best-performing algorithm is the one giving the highest polygenic risk scores (B) Distribution of polygenic risk scores for coronary artery disease (CAD) reported by Khera et al (2018) *Nat Genet* 50:1219–1224; PMID 30104762 with permission from Nature Publishing Group.

Initial efforts were hampered by the small sizes of GWA studies, limited computational methods for predicting genome-wide polygenic risk scores, and a lack of large datasets needed to validate and test the scoring system. However, large datasets from more recent GWA meta-analyses have allowed much greater precision in estimating the impact of individual variants on disease risk. And recently developed large population biobanks, such as the UK Biobank (which has medical and genetic data on 500 000 volunteers), have provided very large datasets for validating and testing the algorithms.

Enthusiasm for polygenic risk scores initially soared after an influential study by Khera et al. in 2018. They reported the development of useful algorithm predictors for each of coronary artery disease, atrial fibrillation, type 2 diabetes, inflammatory bowel disease and breast cancer. In the case of coronary artery disease (CAD), for example, a GWA meta-analysis involving 184 305 CAD cases and controls was the starting point, and 31 algorithms were developed as polygenic predictors of disease risk using the GWA data. The individual predictors were run on a sample of data from 120 281 participants in the UK Biobank and examined to see how well they could detect those participants that had been diagnosed with CAD. The best predictor had an AUC of 0.81 (referenced against a total of over 6.6 million SNP variants). When then used to compute polygenic risk scores for a separate second group of 288 978 UK Biobank participants, the best predictor performed equally well (AUC of 0.81). The testing with this predictor found that 8 % of the population had a genetic predisposition that conferred a threefold or greater risk for CAD, and the prevalence of CAD rose sharply in the highest polygenic score percentiles (see Figure 8.17B).

More recently, however, the usefulness and clinical utility of polygenic risk scores has been questioned. We come back to consider this point when we cover genetic testing in

Chapter 11.

8.3 ASPECTS OF THE GENETIC ARCHITECTURE OF COMPLEX DISEASE AND THE CONTRIBUTIONS OF ENVIRONMENTAL AND EPIGENETIC FACTORS

At the outset of the third decade of the twenty-first century, the information from studies on most common genetic diseases has yet to make a big impact on clinical practice. The medical specialty that has benefited most is oncology: genomic techniques—notably genome-wide sequencing of tumor—have long been deployed in cancer studies. We consider cancers separately in Chapter 10.

For common genetic diseases, GWA studies had been launched in the hope of delivering two major benefits: greater understanding of the molecular basis of disease; and the prospect of developing new drugs and treatments. Many of these diseases have been expected to be collections of different but related diseases; knowing the major genetic determinants might permit disease *stratifica- tion* into disease subtypes to allow more targeted treatments and more efficient clinical management. Cancer genetics has led the way here, and we consider in Chapter 10 how genetic investigations are stratifying cancers into multiple different subtypes.

Novel treatments for a complex disease may also be designed after identifying a *protective factor*, a genetic variant that is negatively correlated with disease. Finally, as new genetic susceptibility factors are identified, there is the prospect of novel *biomarkers*, that is, biological molecules that can be objectively measured and evaluated as indicators of different stages of the disease process; they can be of help in assessing the efficacy of drugs or other new treatments.

As detailed above, the path to understanding the molecular pathologies of complex diseases has not, however, been smooth. Many common genetic diseases are truly polygenic —small contributions can be made by each of a host of different genes that might be expressed slightly differently from normal. We now appreciate that lying between the rare, highly penetrant variants underlying Mendelian disease and the common variants of weak effect, is a much smaller group than initially expected, of low-frequency variants with intermediate effects **(Figure 8.18)**.

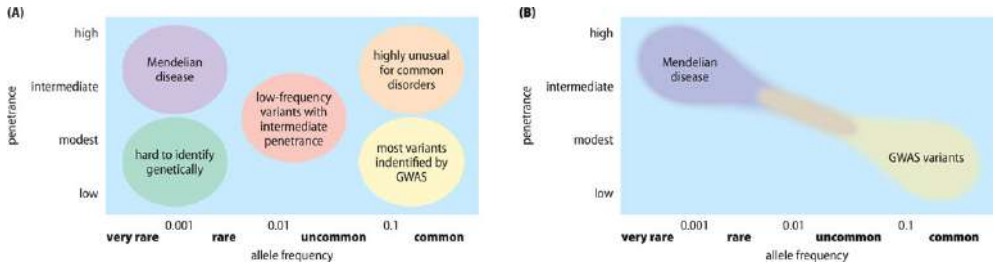


Figure 8.18 Changing views of the roles of genetic factors in determining phenotypes. (A) The hope in the early days of GWA studies. (B) A view after 10 years of GWA studies. (Part A reproduced from McCarthy MI et al. (2008) *Nat Rev Genet* 9:356–369; PMID 18398418, with permission from Nature Publishing Group.)

Because GWA studies rely on linkage disequilibrium to commonly occurring SNP variants, they cannot identify rare variants that may contribute to a complex disease. Rare variants of strong effect may be uncovered in Mendelian subsets, and rare variants of weaker effects might be uncovered through genome-wide sequencing, especially coding DNA variants, or by using micro-array chips designed to identify copy number variants. In this section we focus primarily on how genetic studies have illuminated the molecular basis of common genetic diseases. The genetic studies have involved not just studies of the nuclear genome and SNPs, but also studies of mtDNA variants and of copy number variants.

The genetic architecture of common genetic diseases is a large subject and so we choose to provide select examples to illustrate various principles as follows:

- an outstanding success story: revealing major disease pathways in inflammatory bowel disease (profiled in **Clinical Box 10**)
- how genetic variants underlying Mendelian subsets relate to those in sporadic forms of the same disease: the examples of Alzheimer and Parkinson disease
- the importance of immune system pathways in complex disease
- unexpected linkages between pathogenetic pathways in different diseases and hybrid roles for individual variants as risk factors for some diseases and protective factors for others.

We finish the chapter by taking a look at how non-genetic factors contribute to complex diseases, and examining the role of environmental factors in complex disease and how epigenetic chromatin modifications might be involved.

CLINICAL BOX 10 ILLUMINATING THE PATHOGENESIS OF INFLAMMATORY BOWEL DISEASE (IBD) USING GENOMICS

Inflammatory bowel diseases are characterized by a chronic relapsing intestinal inflammation. Two major subtypes have been distinguished: Crohn's disease, which can occur anywhere along the gastrointestinal tract and affects the entire bowel wall; and ulcerative colitis, which is restricted to the epithelial lining of the colon and rectum. Disease results from abnormal immune responses to commensal organisms, the intestinal microbiota. Heritability is high: estimates from pooled twin studies are 0.75 for Crohn's disease and 0.67 for ulcerative colitis.

Before GWA studies were carried out, almost nothing was known about the genes involved in pathogenetic pathways leading to IBD. One of the very few clues came out of

linkage analyses that ultimately led to identifying the *NOD2* gene as a novel susceptibility factor for Crohn's disease (described in Figure 8.11 and associated text). But right from the outset, GWA studies quickly identified many disease associations and by 2017 they had delivered over 200 risk loci for IBD. The associated genes work in a variety of biological processes (see **Figure 1**), and mostly participate in biological pathways shared by the two disease subtypes.

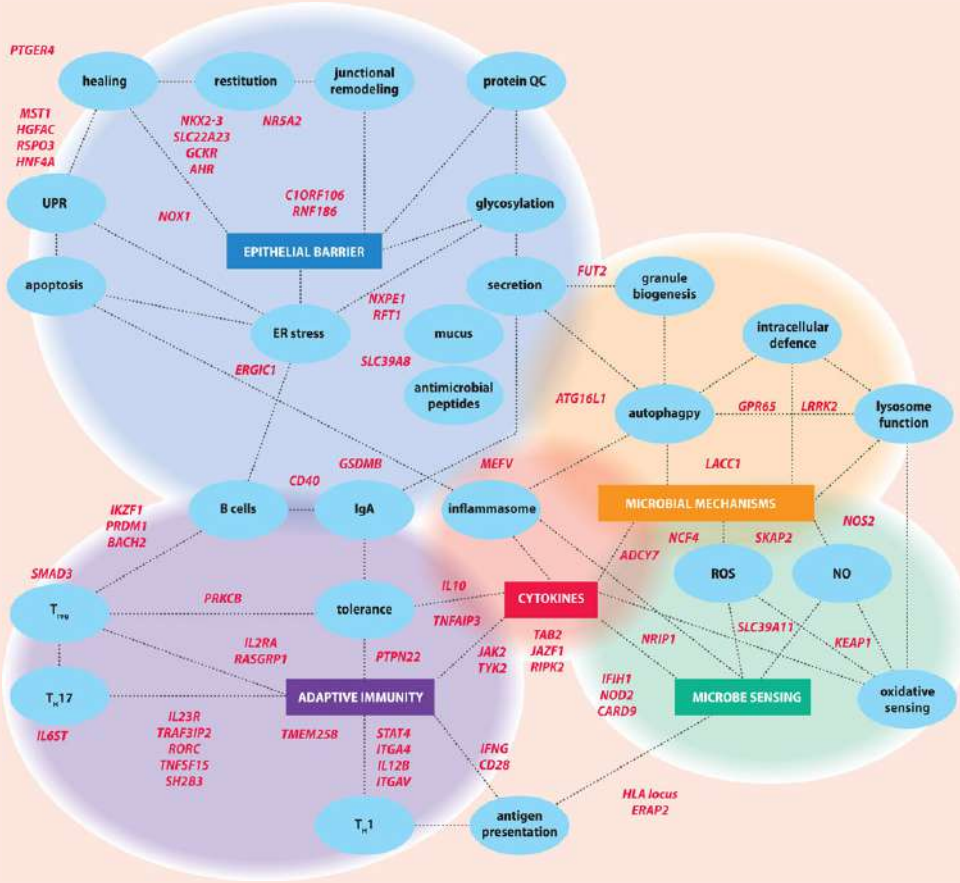


Figure 1 IBD genes and pathways controlling mucosal immunity. IBD risk genes regulate a complex network of interconnected functional pathways. IBD genes (red text) have been implicated in key biological functions (gray circles) that are controlled by interconnected molecular pathways (coloured rectangles). Lines connecting nodes reflect overlapping molecular regulation by common genes. Several IBD risk genes regulate several distinct biological functions depending on their cell type-specific activities. (Reproduced from Graham DB & Xavier RJ [2020] *Nature* 578:527–539; PMID 32103191 with permission from Nature Publishing Group.)

The GWA findings caused a substantial rethink about the pathogenesis of IBD. The importance of some pathways came as a surprise. The GWA risk variants implicated, for example, as many as five genes with a role in autophagy in Crohn's disease (a lysosomal degradation pathway that naturally disposes of worn-out intracellular organelles and very large protein aggregates). The autophagy machinery is now known to interact with many

different stress response pathways in cells, including those involved in controlling immune responses and inflammation.

Another striking—and unexpected—finding was the important role of interleukin-23 (IL-23) pathways in both subtypes of IBD. Tissue injury in these conditions had once been thought to be primarily mediated by classical helper T-cell populations. However, GWA studies have clearly implicated IL-23 and activation of Th17 (a recently discovered subpopulation of helper T cells), resulting in IL-17 production and chronic inflammation. These findings prompted clinical trials using monoclonal antibodies against IL-23, one of which, ustekinumab, has recently been approved for the treatment of both Crohn's and ulcerative colitis.

Common neurodegenerative disease: from monogenic to polygenic disease

Mendelian subsets of common genetic disease are often infrequent, being rare or absent from schizophrenia, bipolar affective disorder, asthma, and stroke. For some complex diseases, however, they are quite common, accounting for ~5 % of all cases of breast cancer, colorectal cancer, and also prostate cancer (where some families show a shared origin with breast and ovarian cancer). In between are disorders where Mendelian subsets are infrequent but significant, accounting for ~1 % of all cases, as for other cancers and coronary heart disease (familial hypercholesterolemia).

Here, we consider the two most common neurodegenerative diseases, Alzheimer disease and Parkinson disease, and the relationship between Mendelian subsets and sporadic disease. In Alzheimer disease, Mendelian subsets are autosomal dominant and rare, accounting for <1 % of all cases, but they are more common in Parkinson disease, accounting for ~5 % of all cases, and comprise both autosomal dominant and autosomal recessive forms.

Linkage analyses provided the first genes to be implicated in these two diseases. Standard linkage analyses were employed for autosomal dominant pedigrees; for autosomal recessive cases autozygosity (homozyogosity) mapping was initially used, but thereafter candidate gene approaches or exome sequencing approaches were employed. A list of gene loci implicated by subsequently identifying disease-specific mutations is given in **Table 8.10**.

TABLE 8.10 GENES UNDERLYING ALZHEIMER (A.D.) AND PARKINSON DISEASE/PARKINSONISM (PARK) IN MENDELIAN SUBSETS

Gene loci (protein product)	Familial disease*	AD/AR	Onset**	OMIM

Gene loci (protein product)	Familial disease*	AD/AR	Onset**	OMIM
APP (amyloid precursor protein);	A.D. type 1	AD	Early	104300
PSEN1 (presenilin1)	A.D. type 3	AD	Early	607822
PSEN2(presenilin2)	A.D. type 4	AD	Early	606889
SNCA (synuclein alpha)	PARK1 & PARK4***	AD	Early/Juvenile	168601/605543
PRKN (parkin)	PARK2	AR	Juvenile	600116
PINK1 (PTEN-induced kinase 1)	PARK6,	AR	Early	605909
PARK7 (DJ1)	PARK7	AR	Early	606324
LRRK2 (leucine rich repeat kinase 2)	PARK8	AD	Late	607060
HTRA2 (HtrA serine peptidase)	PARK13	AD	Late	610297
PLA2G6 (phospholipase A2 group VI	PARK14****	AR	Juvenile	612953
FBXO7 (F-box protein 7)	PARK15*****	AD	Early	260300
VPS35 (vacuolar protein sorting 35)	PARK17	AD	Late	614203
EIF4G1 (Euk. translation initiation factor4?1)	PARK18	AD	Late	614251
DNAJC6 (DnaJ Hsp40 family member C6)	PARK19	AR	Juvenile	615528
SYNJ1 (synaptojanin 1)	PARK20	AR	Early	615530

AD, autosomal dominant. AR, autosomal recessive. OMIM, the Online Mendelian Inheritance in Man database at https://www.omim.org

* Familial disease simply means two or more affected individuals in a family.

** Juvenile onset, average age <40 yrs; early onset, average age <50 yrs; late onset, average age similar to, or slightly less than that for sporadic patients.

*** The *SNCA* gene is hetrozygously triplicated in PARK4, but has missense mutations in PARK1.

**** Also known as adult-onset dystonia Parkinsonism.

***** Also known as Parkinsonian-pyramidal syndrome.

Naturally, heritability is high in Mendelian subsets of both diseases (the causal variants are highly penetrant). In the common polygenic disorders, however, there is a significant difference: common Alzheimer disease is estimated to have a quite high heritability of 0.6 to

0.8, but common Parkinson disease has a low heritability score, around 0.35 to 0.4, and environmental factors are thought to play important roles in Parkinson disease.

The connection between monogenic subsets and sporadic disease

In addition to the gene loci implicated from Mendelian subsets listed in Table 8.10, GWA studies have recently had some successes in Alzheimer and Parkinson disease. The study reported by Schwartzentruber et al. in 2021 (PMID 33589840) describes a GWA meta-analysis of Alzheimer disease that identified 37 risk loci, and a large meta-analysis of Parkinson disease reported 90 significant GWA signals that, however, are almost all located outside coding DNA and explain at most only about one-third of the heritable risk.

The pathology in the monogenic subsets generally mirrors that of sporadic Alzheimer and Parkinson disease, suggesting similar pathways are involved in pathogenesis, even if the monogenic diseases can be more severe, with earlier onset being very frequent. But there may be some phenotype variation as in Parkinson disease types 14 and 15 (Table 8.10).

A series of questions present themselves. Why should some gene loci be implicated in monogenic disease? And can individual gene loci be involved in both monogenetic and sporadic disease? A gene implicated in a monogenic subset must be displaying a rare variant of strong effect, one that has high penetrance. That type of variant cannot be associated with the common sporadic forms of disease, but it does not mean that other variants of the same gene locus are not involved in common disease.

An example of a gene locus involved in both a monogenic subset of Parkinson disease, and also in a common sporadic form, is the *LRRK2* gene. Autosomal dominant Parkinson disease type 8 is due to certain pathogenic missense mutations in *LRRK2*, such as the p.R1441C substitution and p.G2019S substitution, both reported to increase the kinase activity of the LRRK2 protein. (The G2019S mutation is much the more common of the two mutations and is not quite so penetrant.) Common *LRRK2* variants, however, can also act as susceptibility factors for sporadic Parkinson disease, notably those producing the p.R1682P and p.G2385R protein variants. Each of them doubles disease risk (each being found in 3 % to 4 % of healthy individuals but in 6 % to 8 % of individuals with Parkinson disease).

In the final subsection we consider where genes causing autosomal dominant Alzheimer disease act within the known Alzheimer pathogenetic pathways, but first we take a look at the special case of *APOE* in Alzheimer disease.

APOE-ε4, the most significant risk variant in Alzheimer disease

The human *APOE* gene is the major susceptibility locus for Alzheimer disease. The common *APOE*-ε4 allele is such a potent risk factor that in the late 1980s an affecteds-only

nonparametric linkage analysis to look for genes underlying late-onset Alzheimer disease was able to map the disease in some late-onset pedigrees to 19q13, the same location as the previously mapped *APOE* gene. *APOE* makes apolipoprotein E, a key component of lipoprotein complexes that direct the transport and delivery of lipids from one tissue cell type to another. It is produced primarily in the liver, and then in the brain where it has long been known to be a component of the senile plaques characteristic of Alzheimer disease. Differences were found between apoE isoforms in binding to amyloid b (Aβ) *in vitro*, and subsequent candidate gene association analyses confirmed the importance of the *APOE*-e4 allele as a major risk factor in Alzheimer disease.

Apolipoprotein E is a member of the vertebrate family of apolipoproteins, but humans are unique in having a functionally polymorphic apolipoprotein E. As a result of allelic variation, three human apoE protein variants can be produced. The variants show differences at two amino acid positions, and the three different variants confer quite different risk of Alzheimer disease (**Figure 8.19**).

Figure 8.19 Allelic variation at the human *APOE* locus confers significant differences in Alzheimer disease risk. The three human *APOE* alleles, *APOE*ε2, APOE*ε3* and *APOE*ε4*, make proteins—apoE2, apoE3, and apoE4, respectively—that vary at just two out of the 299 amino acid positions. Compared with people homozygous for *APOE*ε3*, the most common allele, people with one copy of the *APOE*ε4* allele have a roughly threefold greater risk of Alzheimer disease, and people with two *APOE*ε4* copies have a roughly fifteenfold increased risk. *APOE*ε2*, by contrast, is a protective allele, conferring a reduced risk compared to *APOE*ε3*.

The *APOE*e4* allele also predisposes to cardiovascular disease as well as to Alzheimer disease (and so will have an effect on reproductive rates). That begs the question of why this allele is so common. Evolutionary studies indicate that *APOE*e4* is the ancestral allele (the monomorphic chimp and gorilla apoE proteins both have arginine residues at positions 112 and 158). *APOE*e4* is thought to have been selectively advantageous to early humans who had a low-calorie, low-fat diet. Over time, however, it has increasingly been replaced by the *APOE*ε 3* allele, which offers the advantage of decreased cholesterol metabolism (reducing the risk of cardiovascular disease).

Rare variants and common susceptibility factors in Alzheimer disease

Rare variants of large effect underlie monogenic forms of Alzheimer disease, but common susceptibility factors are key in sporadic cases. The early-onset and late-onset forms have the same post-mortem brain pathology—abundant extra-cellular plaques, largely composed of amyloid-β (Aβ) peptides of slightly different sizes, and intracellular neurofibrillary tangles mostly made of tau protein. Based on the similar brain pathologies, it had long been supposed that the rare large-effect variants and common susceptibility factors work in common pathways. Molecular studies have confirmed that, as described below.

Amyloid-β (Aβ) is now known to be a central focus of the disease pathways. Aβ peptides are formed by cleavage of the 770-residue transmembrane amyloid-b precursor protein (APP), a neuronal receptor involved in different neuronal functions (including neuronal adhesion and the formation and growth of axons). Aβ peptides are known to be metal chelators, binding to metal ions such as copper, zinc, and iron, and reducing them; they also seem to have antimicrobial function. The Aβ peptides are thought to be the causative agent in Alzheimer disease, partly on the basis of the pathology and on the observation that Aβ is prone to aggregation in the same way as prions (as detailed in Clinical Box 8 on page 229–30), and partly on genetic analyses.

Standard linkage studies of autosomal dominant early-onset Alzheimer disease have identified three causative genes: the *APP* gene, which produces APP, and *PSEN1* and *PSEN2,* which are both involved in processing APP to make Aβ. The APP processing reaction requires sequential cleavage by two endoproteinases: first, a b-secretase (also called BACE1) cuts off most of the large N-terminal extracellular portion of APP; then a multisubunit g-secretase cleaves the trans-membrane segment. The catalytic subunit of g-secretase is a presenilin protein, either presenilin-1 or presenilin-2 (encoded by *PSEN1* and *PSEN2,* respectively).

Cleavage by g-secretase occurs at alternative single locations to generate a series of Aβ isoforms of different lengths (**Figure 8.20A**). The $A\beta_{42}$ isoform (42 residues long) is thought to be the greatest contributor to pathogenesis (it is more prone to forming amyloid aggregates) but is not normally produced in large quantities, unlike the predominant $A\beta_{40}$ isoform.

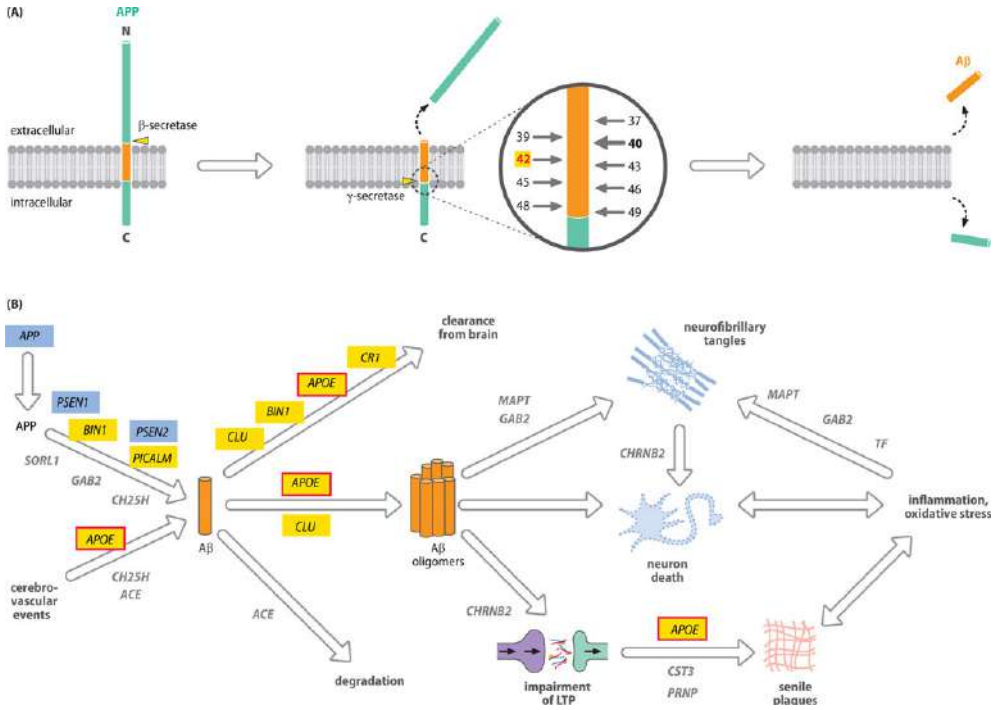


Figure 8.20 Biological pathways in Alzheimer pathogenesis. (A) Production of amyloid-b. The 770-amino acid amyloid precursor protein APP is first cleaved by b-secretase, releasing most of the large extracellular region. Subsequently, the membrane-bound g-secretase cleaves at position 714 or 715, initially generating an amyloid-b (Aβ) peptide 48 or 49 bp long (orange rectangle). It can then go on to trim three nucleotides at a time, generating a set of isoforms of different lengths, from 37 to 49 amino acids long. Of these, Aβ40 is the most frequent isoform, but Aβ42 is especially prone to aggregation. (B) Early-onset Alzheimer disease genes and common late-onset Alzheimer susceptibility factors belong to common pathways. Gene symbols highlighted in blue at the top left are early-onset disease genes in which causal variants are highly penetrant. Some GWA loci are shown (highlighted in yellow); they are common disease susceptibility factors, mostly with generally modest or weak effects. However, *APOE* (shown with a distinguishing red border) has comparatively strong effects and plays key roles in multiple pathways. Other GWA loci shown here are: *BINI,* bridging integrator 1; *CLU,* clusterin; *CR1,* complement component 3b/4b receptor 1; *PICALM,* phosphatidylinositol binding clathrin assembly protein; and *CD33*. Gene symbols given in pale gray were implicated by functional studies. *SORL1* has also been implicated by DNA sequencing in some cases of early-onset Alzheimer. LTP, long-term potentiation. (Adapted from Bertram L & Tanzi RE [2008] *Nat Rev Neurosci* 9:768–778; PMID 22482448. With permission from Macmillan Publishers Ltd.)

Aβ metabolism involves a balance between the production from APP and its removal, either by enzymatic degradation (proteolysis) or by receptor-mediated transport out of the brain via the blood-brain barrier (clearance). Pathogenesis results from an increase in the amount of Aβ or the amount of Aβ42 relative to

Aβ40, and soluble Aβ oligomers may have a primary contribution (in addition to affecting synaptic transmission—by impairing long-term potentiation). They may exert some of their effects by regulating the production and phosphorylation of tau protein to induce the

formation of neurofibrillary tangles that can cause neurons to die (Figure 8.20B). Aβ oligomers can further aggregate into fibrils that end up in extracellular senile plaques that can provoke inflammation responses that can further contribute to pathogenesis.

GWA studies have identified additional variants, some of which have been well replicated and are considered established susceptibility factors. Like *APOE*, they have been implicated in pathways involving Aβ, but principally in the production of Aβ and its clearance from the brain (see Figure 8.20B). However, several of the genes have a role in inflammation *(CR1* and *CLU*) or the innate immune response (*CD33;* not shown). None of the newly implicated variants in late-onset susceptibility have strong effects (typically odds ratios of 1.15 or 1.10). But as the biological pathways in disease are mapped, new targets may become available for drug therapy.

The importance of immune system pathways in common genetic disease

We have long been aware of the importance of immune system pathways in a subset of complex disease: autoimmune diseases. Variants in the HLA complex at 6p21.3 are the strongest genetic risk factors for diseases such as rheumatoid arthritis, type 1 diabetes, systemic lupus erythematosus, multiple sclerosis, celiac disease, myasthenia gravis, Graves' disease, psoriasis, and so on. Earlier in this section, the importance of immune system pathways in the pathogenesis of inflammatory bowel disease could be seen in the bottom half of Figure 1 in Clinical Box 10 above. Chronic low-grade inflammation is also present in type 2 diabetes, a metabolic disorder (as well as in the autoimmune type 1 diabetes). Activated innate immune cells accumulate in metabolic tissues with the release of inflammatory mediators.

Although the eye and the brain have been considered as immune-privileged organs—ones where foreign tissue grafts can survive for extended time periods while similar grafts placed at most other sites in the body are acutely rejected—GWA studies have revealed the importance of certain immune system pathways, notably innate immunity, in common genetic disorders of the eye and brain. Take the most common cause of blindness in developed countries as an example: age-related macular degeneration. In this condition, the macula, a small patch of the retina responsible for central vision, is affected. A very early GWA study, using a small number of cases and controls only, identified the *CFH* (complement factor H) gene as a major risk factor in this disorder—a case of a common variant with an unusually strong effect. Follow-up GWA studies also identified three other complement genes as prominent susceptibility factors—*C2* and *CFB* (neighboring genes in the class III HLA region, encoding complement C2 and complement factor B, respectively)—and the *C3* gene at 19q13.

Immune cells actively contribute to homeostatic processes in the nervous system, and include microglia, the brain's primary resident immune cells, mast cells (in the parenchyma) and even small populations of T and B cells in the developing brain. It may not be altogether surprising, then, that immunity pathways are important in common genetic disorders

affecting the brain. As an example of neurodegenerative disease, consider Alzheimer disease. As mentioned above, GWA studies have implicated several genes with a role in inflammation—*such as the CR1* (complement C3b/C4b receptor 1) and *CLU* (clusterin) genes (see Figure 8.20)—or in the innate immune response, such as *CD33*. The accumulation of cerebral b-amyloid plaques is thought to result from imbalanced production and removal of amyloid-b peptides, arising from innate immune cells losing the ability to restrict their accumulation.

The immune system plays an important role in neurodevelopment, regulating neuronal proliferation, synapse formation, and plasticity, as well as removing apoptotic neurons. As an example of a neurodevelopmental disorder, consider autism spectrum disorder (ASD) where immune dysfunction in ASD has been repeatedly described, with symptoms including neuroinflammation, increased T cell responses, autoantibodies and enhanced innate NK cell and monocyte immune responses. Not unexpectedly, therefore, innate immune response genes are dysregulated in autism (as revealed by transcriptome analysis—interested readers can find an example at PMID 25494366).

Finally, taking schizophrenia as an example of a psychiatric disorder, GWA studies have identified more than 30 susceptibility factors known to have an immune system function or to be expressed in T or B lymphocytes (PMID 29701842). The strongest genetic risk factor is genetic variation within the major histocompatibility complex. Classical class I or class II HLA genes are not involved but, yet again, a complement gene is: an increased number of complement *C4* genes in the class III region of the HLA complex is a very strong disease risk factor in schizophrenia. Because of evolutionarily recent tandem duplication of an ~30 kb segment of DNA, there are two slightly different complement *C4* genes (see **Figure 8.21A**).

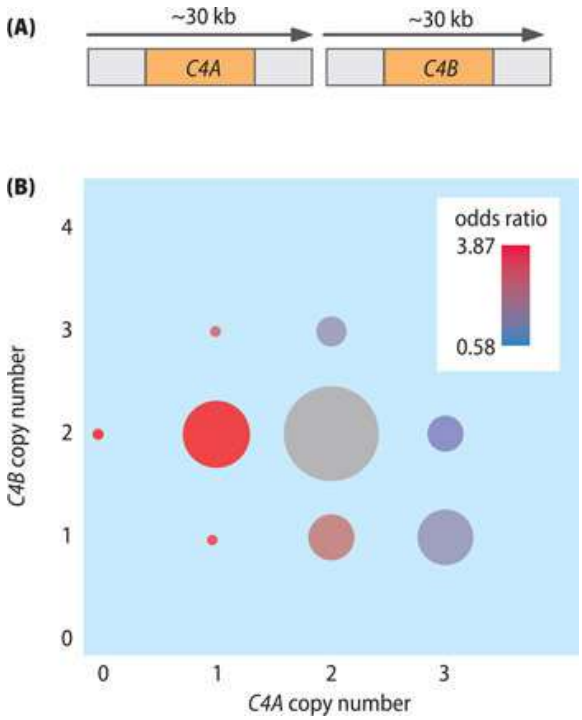


Figure 8.21 Susceptibility to, and protection against, systemic lupus erythematosus (SLE) correlates with complement *C4* gene copy number. (A) The standard (= most common) haplotype has a tandem duplication of ~30 kb with two slightly different *C4* genes, *C4A* and *C4B*. (B) Decreased *C4* copy number is a risk factor for SLE, notably when there is a single *C4A* allele (high odds ratio); increased *C4* copy number, notably three *C4A* copies protects against SLE (low odds ratio). The area of each circle is proportional to the number of individuals with that number of *C4A* and *C4B* genes. Panel B reproduced with permission from Kamitaki N et al. (2020) *Nature* 582: 577–581; PMID 32499649

In addition to the most common haplotype (which has two C4 genes), unequal crossover can result in haplotypes with one, three, or sometimes four *C4* genes, and excess *C4* genes predispose to schizophrenia. Complement C4 proteins are found in neuronal synapses, dendrites, axons, and cell bodies, and mouse studies suggest a role for complement C4 in synaptic pruning. Producing an excess of complement C4 is thought to cause the reduction in the number of synapses seen in the brains of schizophrenia individuals. A strong risk factor for schizophrenia, increased complement *C4* copy number is simultaneously a protective factor for some other diseases, as described in the next section.

The importance of protective factors and how a susceptibility factor for one complex disease may be a protective factor for another disease

Unravelling the molecular pathology of complex reveals connections between molecular components and biological pathways in different diseases. For example, the common R620W variant of the PTPN22 protein is known to modify disease risk in several autoimmune disorders. But what is now emerging are unexpected links between rather

different diseases. According to the extent to which they share GWA variant profiles, heatmaps can be generated to compare the genetic profiles of different diseases—interested readers can find an example in Figure 1 of PMID 20041220.

As well as susceptibility factors, which confer increased risk of disease, genetic investigations are identifying a series of protective factors that reduce disease risk. In most cases identified protective alleles are genetic variants that result from point mutation. They often produce a changed protein that works in a different way, but some are nonsense mutations or splice variants—see **Table 8.11** for examples.

TABLE 8.11 EXAMPLES OF PROTECTIVE VARIANTS OR ALLELES THAT REDUCE THE RISK FOR A COMMON DISEASE

Protective variant	Disease	Comments
POINT MUTATIONS		
*APOE*ε2*	Alzheimer disease	common allele (has cysteines at positions 112 and 158—see Figure 8.19)
*APP*A673T*		inhibits cleavage of APP, reducing production of amyloid-β
Blood group O	Coronary heart disease	AB blood group is a significant risk factor
*CARD9*IVS11 + 1G>C*	Crohn's disease	a rare splice variant, but highly protective (odds ratio 0.29)
*PTPN22*R620W*		common allele; simultaneously a strong risk factor for both type 1 diabetes and rheumatoid arthritis
*HLA-DRB1*1301*	Rheumatoid arthritis	Affords protection in the 70% of affected cases who have anti-citrullinated protein antibodies
*PCSK9*C679X*	Coronary artery disease	*PCSK9* is important in cholesterol homeostasis, and inactivation reduces lipid levels; the C679X allele has a frequency of 1.8% in the US black population
COPY NUMBER CHANGES		
Complement *C4A* gene increased copy number	SLE (lupus), Sjogren syndrome	see Figure 8.21 for SLE; simultaneously a risk factor for schizophrenia

Non-HLA genes: *APOE,* apolipoprotein E; *APP,* amyloid protein precursor; *CARD9,* caspase recruitment domain family, member 9; *CCR5,* chemokine (C-C motif) receptor 5; *PCSK9,* proprotein convertase subtilisin/kexin type 9; *PTPN22,* protein tyrosine phosphatase, non-receptor type 22.

Protective variant	Disease	Comments
Complement *C4A* gene decreased copy number	schizophrenia	see text; simultaneously a risk factor for SLE and Sjogren syndrome

Non-HLA genes: *APOE,* apolipoprotein E; *APP,* amyloid protein precursor; *CARD9,* caspase recruitment domain family, member 9; *CCR5,* chemokine (C-C motif) receptor 5; *PCSK9,* proprotein convertase subtilisin/kexin type 9; *PTPN22,* protein tyrosine phosphatase, non-receptor type 22.

The lower part of Table 8.11 lists examples of DNA variants that act as protective factors through gene dosage, that is alteration in the copy number of a gene. Increasing the number of complement C4 genes, especially the *C4A* gene, provides protection against lupus; conversely, a reduced number of complement C4 genes is a risk factor for lupus—see Figure 8.21. The reverse is true for schizophrenia, as noted above.

A strong risk factor for a common genetic disease may also be a protective factor for an infectious disease. One such example is the common *FUT2* non-secretor allele, a nonsense mutation in the gene that makes a(1,2)-fucosyltransferase. This enzyme completes the synthesis of H antigens, precursors of the ABO histo-blood group antigens found on cells in body fluids and on the surface of the intestinal mucosa. Homozygotes for the non-secretor allele fail to present ABO antigens in secretions and in the intestinal mucosa and have an increased risk of Crohn's disease and type 1 diabetes (most probably because of alterations to the diverse microorganisms resident in the gut). But the same individuals are strongly resistant to some strains of norovirus, the most common cause of non-bacterial gastroenteritis. A balance between acting as a risk factor for one common condition and acting as a protective factor for a different condition may explain why certain genetic variants associated with risk of a common genetic disease have reached relatively high frequencies.

Gene–environment interactions in complex disease

Environmental factors are clearly important in cancers and infectious disease where some associations, such as *Helicobacter pylori* infection and ulcer formation, have only recently become evident. But they are increasingly being recognized to be important in complex diseases outside those two categories (as well as in some monogenic disorders); see **Table 8.12**. That should not be surprising because monozygotic (identical) twins are often seen to be discordant for complex diseases (as shown previously in Table 8.4). Identical twins arise from the same zygote and might be expected to have identical DNA profiles. If one twin develops a complex disease, such as Crohn's disease, but the other lives a long healthy life,

factors other than DNA can be expected to be important (although post-zygotic mutations could also have a role in some cases).

TABLE 8.12 EXAMPLES OF DIFFERENT TYPES OF ENVIRONMENTAL FACTORS THAT CONTRIBUTE TO COMMON NON-INFECTIOUS AND NON-CANCEROUS DISEASE

Environmental	Examples
Teratogens and abnormal metabolite levels in uterine environment	low folic acid increases the risk of neural tube defects such as spina bifida
Unbalanced diets	over-consumption and excess of fatty foods can predispose to type 2 diabetes
Smoking	increased risk for disorders such as coronary artery disease, Crohn's disease*, and aging-related macular degeneration
Commensal microorganisms	gut microbiota in inflammatory bowel disease, type 2 diabetes

* But not ulcerative colitis-if anything, smoking protects against ulcerative colitis.

Striking evidence for the importance of environmental factors comes from increased risk for a specific disease that often befalls migrants who have moved from a community with a low general risk of that disease to join a society in which the disease is much more prevalent. In addition, as populations across the globe change their eating habits and lifestyles, there is a relentless rise in the frequency of conditions such as obesity and type 2 diabetes.

Gene–environment interactions are also important in the sense that they can make it difficult to detect a genetic (or environmental) effect if they are not identified and controlled for. That can lead to inconsistent disease associations when populations are variably exposed to certain environmental factors that modify the effect of a given genetic variant (**Figure 8.22**). Understanding gene-environment interactions can therefore allow us to develop protective strategies for complex diseases: by seeking to minimize exposure to an environmental factor, the harmful effect of a genetic susceptibility factor can be minimized.

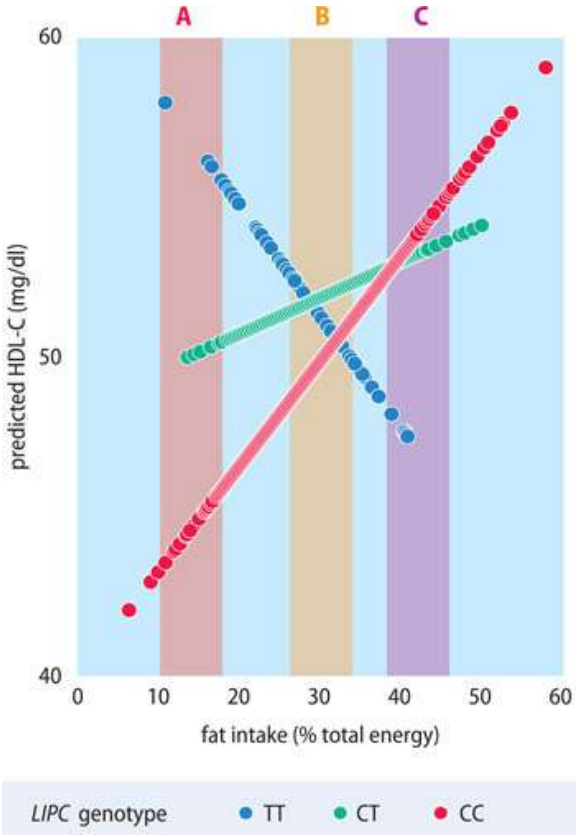


Figure 8.22 The importance of gene–environment interactions—an example. According to different total levels of dietary fat intake, variations in predicted values of high-density lipoprotein cholesterol (HDL-C) are shown for three genotypes at the –514(C/T) polymorphism (rs1800588) in the *LIPC* hepatic lipase locus. Low fat intake (band A) combined with the TT genotype (homozygous for the T allele) results in the highest HDL-C level. For a moderate fat intake (band B), there is no relationship between genotype and HDL-C level. For a high fat intake (band C), the TT genotype has the lowest HDL-C level. Gene–environment interactions are therefore important in identifying genetic and environmental determinants of medically relevant phenotypes such as HDL-C levels; depending on the dietary fat intake, one could variously conclude that the TT genotype produces high HDL-C levels (band A) or low HDL-C levels (band C), or that it is not associated with HDL-C levels at all (band B). (Adapted from Manolio TA et al. [2006] *Nat Rev Genet* 7:812–820; PMID 16983377. With permission from Macmillan Publishers Ltd.)

A plethora of "environmental factors"

The term *environment* has a multi-layered meaning in this context—it effectively includes any component that alters disease risk without originating from DNA variation in our cells. There are external physical environments. Right from the earliest stages and throughout life, we are exposed to a range of diverse radiation sources, and also to infectious agents that can influence susceptibility to non-infectious diseases. In the womb, we are exposed to a uterine environment and will be variously affected by what our mother consumes during pregnancy.

After birth and throughout life, we ingest, or have surface contact with, a huge range of additional foreign molecules. The molecules that we ingest intentionally (in food and drink, stimulants, and so on) may be considered, in part, lifestyle choices. They, too, along with the amount of physical and mental exercise that we experience and the degree of stress to which we are subjected, are important in disease susceptibility.

Then there is our internal microbiome, the diverse range of microorganisms *(microbiota)* that constitute part of us. Mostly composed of bacteria, our personal microbiomes have 10 times more cells than we have and are mostly located within the gut (the gut microbiome in an average person has possibly 5000 different bacterial species. In addition to being beneficial to us (see above), our microbiomes have a major influence on susceptibility to disease—see above and the review by Virgin & Todd (2011) under Further Reading. Finally, there is the environment inside our body cells and how it links with the extracellular environment. Two important, and interlinked, components here are mtDNA variation (**Box 8.6**) and chromatin modifications, including DNA methylation.

BOX 8.6 MITOCHONDRIAL DNA HAPLOGROUPS AND MITOCHONDRIAL DNA VARIATION IN COMMON DISEASE

Because mitochondria are the power sources of our cells, the performance of cells (notably brain and muscle cells, which have high energy requirements) is hugely dependent on mitochondrial efficiency. In an environment where food (and therefore calories) is plentiful, mitochondria efficiently generate energy to keep cells in optimal condition; severely restricting calorie intake impairs mitochondrial and cell efficiency.

Mitochondria have an important influence on how nuclear genes are expressed, because they make the ATP and acetyl coenzyme A needed for diverse cell signaling pathways and for phosphorylating and acetylating histones in chromatin. They are also the principal generators of reactive oxygen species that damage our cells. Ageing is a major risk factor for most common diseases, and accumulating oxidative damage through a lifetime is a principal contributor to increasing cellular inefficiency. The genetic control of mitochondrial function is mostly specified by nuclear genes, but mitochondrial DNA (mtDNA) is much more susceptible to mutation than is nuclear DNA.

EVOLUTION OF mtDNA HAPLOGROUPS

Because mtDNA is strictly maternally inherited, mtDNA undergoes negligible recombination at the population level, and so SNPs in mtDNA form branches of an evolving phylogenetic tree. The major subdivisions of the world mtDNA phylogeny occurred more than 10 000 years ago and are called mtDNA haplogroups, which developed as humans migrated into new geographic regions, leading to region-specific haplogroup variation (**Figure 1**).

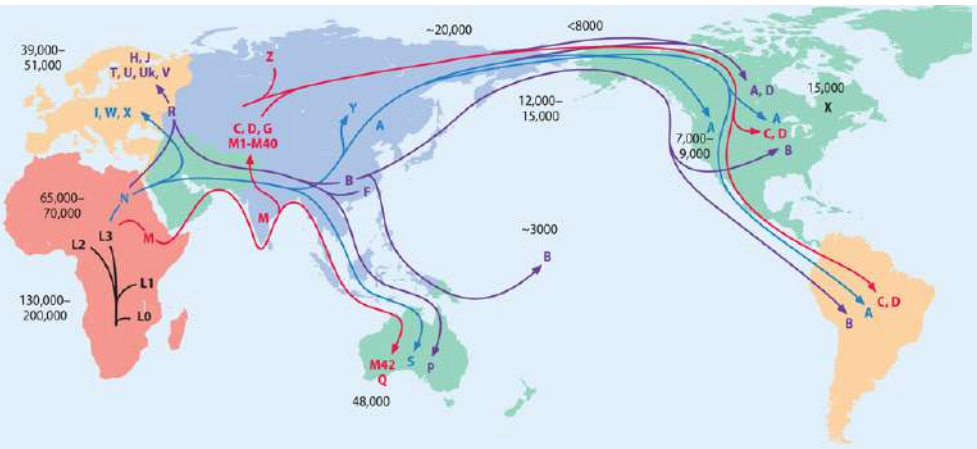


Figure 1 Evolution of mtDNA haplogroups. The estimated mutation rate is 2.2–2.9 % per million years. Time estimates are in years before the present. (From the MITOMAP database at http://www.mitomap.org)

More than 95 % of Europeans belong to one of 10 major haplogroups, namely H, J, T, U, K (a subgroup of U), M, I, V, W, and X; several of these are associated with complex human traits (see below for examples). Each haplogroup defines a group or "clade" of related mt DNAs containing specific sequence variants within the population. Mitochondrial DNA haplogroups influence the assembly and stability of the mitochondrial respiratory chain, the synthesis of respiratory chain proteins, and the propensity to develop intracellular oxygen free radicals, which are implicated in the pathophysiology of several common human diseases.

Although contentious, some evidence suggests that the distribution of human mtDNA haplogroups has been influenced by environmental pressures, including climate. (Nuclear-mitochondrial DNA coevolution has been implicated in climatic evolution in other species.)

MITOCHONDRIAL DNA AND COMMON DISEASE

A variety of rare mitochondrial disorders is known to be due to variants of large effect in single genes in mtDNA, or rare multigenic mtDNA deletions and duplications. Common mtDNA variants with weak effects, notably single nucleotide variants, are known to alter the penetrance of these rare disorders. More recently, diverse association studies have shown that common mtDNA variants also influence susceptibility to a wide range of complex diseases, including neurodegenerative, psychiatric, cardiovascular, and many other diseases—the supplementary table in Gomez-Duran A et al. (2010) *Hum Mol Genet* 19:3343–3353 (PMID 20566709) gives a list of examples.

Certain mtDNA haplogroups have been shown to be associated with disease. The most common mtDNA haplogroup H, found in about 40 % of Europeans, is associated with a more than two-fold increased change in surviving severe infection (sepsis), but subgroups of this haplogroup are emerging as a risk factor for late-onset degenerative diseases,

including those affecting the nervous system. This may suggest that infectious disease has shaped mtDNA evolution in Europe over a relatively short period, increasing the frequency of mtDNA haplogroup H and thereby predisposing modern humans to late-onset common disease.

Despite the frequent implication of mtDNA in different common complex diseases, the precise haplogroup associations are not always consistent. This is partly due to the limited cohort size in some studies, restricting the statistical power. Another issue is the different frequency of mtDNA haplogroups in different ethnic groups. A further confounder is occurrence of the same base substitution on different branches of the phylogenetic tree as a result of recurrent mutation, called mtDNA *homoplasy* (but not to be confused with homoplasmy), which accounts for up to 20 % of genetic variation in Europeans. The frequency of a subhaplotype containing a functional homoplasy can vary in different populations, and the distribution of subhaplogroups also varies in different populations. As a result, the major haplogroups can be associated with the disease in some populations but not in others.

The study of gene–environment (GxE) interactions has traditionally involved case-control studies of candidate genes, but the advent of GWA studies has prompted hypothesis-free genome-wide studies. They require large sample sizes, however—a GxE GWA study needs about four times as many samples as a standard GWA study to detect a main effect of the same magnitude. Various GxE GWA studies have been launched, such as a scan to identify genes conferring susceptibility to air pollution in childhood asthma.

Prospective cohort studies

Case-control studies are the most widely used method of investigating the genetic and environmental basis of complex disease. Cases and controls are typically investigated *retrospectively* (that is, the disease cases have already occurred, and subjects need to be quizzed about previous events such as exposure to environmental factors). As a result, the studies are open to all kinds of bias, for example in the selection of subjects to be studied.

Prospective cohort studies have the big advantage of removing much of the bias by studying individuals over a long timeframe that commences *before* the onset of disease. They involve periodic assessment of subjects, including recording detailed information on them and collecting samples for future laboratory tests. Studies such as these do not select affected individuals, and so need to be very large to ensure that eventually there will be statistically significant numbers of affected individuals.

A leading example of a prospective cohort study is the UK Biobank project. From 2007 to 2010 it recruited 503 000 British people aged between 40 and 69 years and will go on to follow them with periodic testing over a period of 30 years (**Table 8.13**). By comparing

those who remain healthy with those who develop disease within the 30-year timeframe of the study, researchers hope to gain important information on the genetic determinants of a range of common late-onset diseases (including cancers, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression, and forms of dementia). The study will also help measure the extent to which individual diseases have genetic and environmental causes.

TABLE 8.13 COMPONENTS OF THE UK BIOBANK PROSPECTIVE COHORT STUDY

Baseline questionnaire	Baseline physical measurements	Follow and future measures
Sociodemographic	blood pressure	stored blood, urine, saliva
Family history	weight, body impedance	repeat baseline assessment (20000 participants)
Environmental	waist and hip circumference	
Lifestyle	seated and standing heights	access national health records: death, cancer, hospitalizations, primary care
Cognitive function	grip strength	
Food frequency	bone density	
Internet-administered 24-hour dietary questionnaire	mailed triaxial accelerometers enhanced phenotyping (last 100000-150000 participants recruited): hearing, vascular reactivity, visual acuity, refractive error, intraocular pressure, corneal biomechanics, optical coherence tomography, fitness assessment	

From ManolioTAet al. *(2012) Am J Epidemiol* 175:859-866; PMID 22411865. With permission from Oxford University Press.

Epigenetics in complex disease and aging: significance and experimental approaches

How do environmental factors work to have an impact on complex disease? Somehow they must affect how our genes are expressed. They can do that at the DNA level by changing the DNA sequence in our cells (recall the environmental mutagens that we considered in Section 4.1).

Alternatively, if they are infectious agents they can introduce some novel genes or proteins that change how our cells work. Yet another way—and one that is now seen to be very common—is to change the epigenetic settings (the *epi- genome*) in our cells. We

described before how epigenetic effects regulate gene expression (Section 6.2) and how they are important in some monogenic disorders (Section 6.3). And in Chapter 10 we illustrate how epigenetic effects are very important in cancer. Here, we focus on epigenetic effects in other complex diseases.

Unlike the genome, which is very stable, the **epigenome**—effectively the chromatin states across all chromosomes (determined primarily by patterns of cytosine methylation, histone modification, and the positions of nucleosomes)—is comparatively fluid. In response to certain environmental cues (signals), the epigenome can be significantly altered, and that can result in important changes in gene expression.

As described in the second and third subsections below, epigenetic changes occur throughout the life of an individual. They are thought to be important in aging—a frequent risk factor in complex disease—and they can explain, at least in part, why identical twins develop to become different (different post-zygotic mutations arising in the identical twins can also be expected to play a part).

As shown in Figure 6.16 on page 158, early development is a period of rapid changes in the epigenomes of cell, with a global resetting of methylation marks. And at this stage epigenomes can be particularly sensitive to environmental factors. As described below, a popular theory holds that chronic adult diseases originate in early life, and perturbation of epigenetic settings by environmental factors is an attractive explanation. We consider ways in which that can happen below.

Experimental investigations

Because epigenomes are highly variable between different types of cell, analyzing epigenomes is potentially more complicated than genome analysis. As a result, the investigation of epigenetic factors in complex disease has lagged behind genome analysis. In the past few years, however, great strides have been taken in defining epigenetic settings in cells.

Analysis of global patterns of DNA methylation (the "methylome") is comparatively advanced—the positions of 5-methylcytosines have been mapped across the genome to single nucleotide resolution in some cell types. Investigators can now carry out large-scale DNA methylation scans across the genome by using microarrays such as Illumina's Infinium HumanMethylation450 BeadChip (it scans 485 000 cytosine methylation sites distributed across virtually all protein-coding genes with an average of 17 CpG sites per gene region, including CpG sites in the promoter, untranslated sequences, first exon, and elsewhere in the gene body).

Genome-wide investigations to identify the extent of epigenetic contributions to complex disease have been launched for common neurological and autoimmune disorders and some other disorders.

Epigenetic changes during aging

Aging, a very important risk factor for complex disease, is marked by progressive inefficiency in cell, tissue, and organ function. There are inherent limits on the efficiency of cellular processes, including endogenous errors in DNA replication, in DNA repair, and in the regulation of gene expression. As we age, therefore, both genetic and epigenetic changes accumulate progressively, and changes in the genomes and epigenomes of somatic stem cells (that are involved in maintaining tissue homeostasis) may be fundamental in the aging process.

The epigenetic changes that accumulate in our cells can be secondary to genetic changes (mutations in DNA sequences that regulate epigenetic mechanisms) or to inherent errors in the epigenetic regulation machinery. However, they can quite often be induced by environmental factors or as a result of stochastic (chance) factors. Both cytosine methylation and histone modification patterns change with aging. In the former case, for example, there is a progressive loss of cytosine methylation across the genome during aging, but against this general pattern of global hypomethylation (which includes very many methylation sites outside gene regions), hypermethylation occurs at promoters of certain genes.

Epigenetic changes in monozygotic twins

Epigenetic changes may constitute a major reason why monozygotic (identical) twins, who are initially extremely similar in appearance and behavior, go on to develop significant differences in various aspects of the healthy phenotype. And there is quite frequent discordance between identical twins for a variety of complex genetic disorders.

Identical twins derive from a single zygote (the embryo splits at a very early stage in embryonic development), and so they initially have identical genetic profiles (but may accumulate different post-zygotic mutations). Epigenetic differences between identical twins are initially minimal but can begin to occur even in prenatal development. Stochastic factors may be involved, such as different X-chromosome methylation patterns in female identical twins arising from the random choice of which parental X chromosome to inactivate.

Environmental factors can also play a part in prenatal development, because the *in utero* environments can be different. The vast majority of identical twins are located in separate amniotic membranes, and in diamniotic twins there is an increased risk of congenital heart disease that usually affects just one twin. Differences in exposure to postnatal environmental factors most probably contribute to the significant epigenetic differences observed in older monozygotic twins (in both cytosine methylation and histone modification patterns).

The developmental origins of adult health and disease

Pioneering epidemiological studies have established that low birth weight confers an increased risk of developing different common adult diseases, including various cardiovascular diseases, hypertension, and stroke. When significantly fewer nutrients are provided to the fetus during pregnancy, reprogramming in early development seems to cause the fetus to develop a "thrifty phenotype" with a low metabolic rate and reduced pancreatic beta cell mass and islet function. The *thrifty phenotype* is thought to be an adaptation that maximizes the chance of surviving in an adverse environment where calorie intake is restricted, but the altered metabolism is not well adapted to a later life where food is plentiful, increasing the risk of metabolic syndrome (with strong risk determinants for type 2 diabetes, obesity, and hypertension).

The effects of the Dutch *Hongerwinter,* a wartime famine that took place in western parts of the Netherlands for six months in 1944/1945, provide support for the "thrifty phenotype" hypothesis. Women who endured semi-starvation conditions during mid to late gestation gave birth to underweight babies who were then exposed in later life to normal levels of calorie intake, with an increased incidence of common metabolic and cardiovascular diseases. Individuals born to mothers who experienced starvation conditions during early gestation only, so that they were of average weight at birth, had even higher rates of obesity than those who suffered sharply reduced nutrition in mid to late gestation, and they also had an increased risk of schizophrenia. By implication, early gestation appears to be a particularly critical time in which environmental factors can have an influence.

Because other nutritional cues during infancy and childhood were also found to be associated with adverse effects in later life, the "thrifty phenotype" hypothesis has broadened into a more general theory that proposes that a wide range of environmental conditions during embryonic development and early life determine susceptibility to different adult diseases.

Environmental factors seem to have an impact on development so as to increase the risk of disease in later life, but how do they work? The comparative plasticity of epigenomes makes them likely targets of environmental factors, and this is supported by data from experimental models and human studies (notably environmentally induced changes in DNA methylation patterns). Because epigenetic processes, such as DNA methylation and histone modifications, rely on metabolic factors, a differential availability of dietary components can be expected to influence epigenetic mechanisms. For example, methylation of cytosines and histones uses *S*-adenosylmethionine as a methyl donor, and dietary factors, notably folate (vitamin B9), are known to have a key role in the pathway that produces *S*-adenosylmethionine. As described in Section 10.3, cancer studies have also shown a direct link between inflammation (which is often triggered by environmental factors) and epigenetic modification causing altered gene expression.

Transgenerational epigenetic effects

Epigenetic effects are clearly transmitted through mitosis so that chromatin states are heritable through cell generations. For example, when a liver cell divides it gives rise to two liver cells with the same type of epigenome (genome-wide pattern of chromatin states) as the parent cell. But can epigenetic effects be transmitted through meiosis? Might a pattern of increased disease risk deriving from environmentally induced epigenetic modifications be passed on to children so that they, too, have increased disease risk?

Transgenerational epigenetic inheritance is common in plants but rare in animals. In the nematode worm *Caenorhabditis elegans*, experimental manipulation of specific chromatin modifiers in parents can result in an extended lifespan up to the third generation. Suggestive evidence for human transgenerational epigenetic effects comes from certain studies in northern Europe, such as a Swedish study that seems to link the availability of food supply during the early life of the *paternal* grandparents and longevity of the grandchildren, including associations with cardiovascular disease and diabetes. However, a defined mechanism for transgenerational inheritance in humans and animal models is currently missing. Interested readers should consult recent reviews such as those by Grossnicklaus et al. (2013) and Cavalli and Heard (2019) listed under Further Reading.

SUMMARY

- Identifying genes for Mendelian disorders used to rely on first finding a subchromosomal location for the disease gene, usually by genome-wide linkage analyses, but sometimes by looking for disease-associated chromosome breakpoints. Genes would be sought in the target region, and promising candidates tested for evidence of disease-associated mutations.
- Genetic linkage investigates whether alleles at two or more loci co-segregate in families. Two loci located on different chromosomes, or far apart on the same chromosome, are unlinked; alleles at the loci will have a 50 % chance of being inherited together at meiosis. Alleles at linked loci (lying close together on a single chromosome) will often be co-inherited as a haplotype.
- For Mendelian disorders, parametric linkage analyses would be used, ones where the mode of inheritance, disease gene frequency, and penetrance of disease genotypes were inputted into the program. Genome-wide linkage analyses to map a Mendelian disorder require several hundred polymorphic markers from across the genome. If a marker locus is physically close to the disease locus, a marker allele will tend to co-segregate with disease during meiosis.
- The modern alternative for finding Mendelian genes involves whole exome sequencing (sequencing the exons plus immediately flanking intron sequence of protein-coding genes plus miRNA genes).

- A polygenic trait, such as adult height or blood pres sure, shows a continuous range of values and a normal (bell-shaped) distribution within a population. The genetic susceptibility is due to alleles at many loci, each of weak effect.
- In complex (multifactorial) diseases no single gene locus dominates to the same extent as in monogenic conditions. As well as being polygenic, various nongenetic (environmental) factors have important roles in disease.
- DNA variants that cause a Mendelian disorder are rare variants of strong effect. By contrast, many DNA variants involved in pathogenesis of a complex disease are common (occur at high frequency) and are of weak effect. They are susceptibility factors, being found in unaffected individuals but at lower frequencies than in affected people.
- In complex disease, the disease susceptibility needs to cross some high threshold value for a person to be affected. Affected people will have high-risk alleles at multiple susceptibility loci, and so their first-degree relatives will be at higher risk than the general population. Depending on environmental factors, the liability threshold value can change and often shows sex differences.
- The variance of a phenotype is the square of the stan dard deviation; the heritability is the proportion of the variance that is due to genetic factors.
- In diseases with a strong genetic component, siblings of an affected person have a much higher risk of disease than the general population, and monozygotic twins are much more likely to show concordance in disease status than are dizygotic twins.
- Heritability is not a fixed property—for any disease it varies between populations, and can vary within the same population when environmental factors change.
- Except in the case of rare Mendelian subsets, linkage analyses used in complex disease are nonparametric. They test affected relatives only, typically affected sibs, and look for chromosomal segments that they share more often than expected by chance.
- Association studies are superior to linkage analyses in finding susceptibility factors for complex diseases. They test affected individuals (cases) and unrelated controls from the same population to seek *statistical* associations between individual variants and the disease.
- Association studies aim to identify common alleles (on short chromosome segments) that have significantly different frequencies in cases and controls. The associations may arise because many people share a short chromosome segment inherited from a distant common ancestor who carried a susceptibility factor.

- The international HapMap project has defined ances tral chromosome segments in various human populations. The shared ancestral chromosome segments are usually very small (often just a few kilobases).
- Candidate gene association studies test for associa tion between a disease and alleles of a specified gene of interest (often because of a suspected role in the disease).
- Association works over very short ranges only on the DNA, and so in genomewide association (GWA), densely spaced markers are needed (usually with at least 500 000 SNP markers across the genome).
- DNA variants may confer increased disease risk (sus ceptibility factors) or reduced risk (protective factors). A risk factor for one complex disease may sometimes be a protective factor for a different disease.
- Association studies may reveal haplotype blocks har boring a disease susceptibility variant but identifying the pathogenic variant can often be difficult because of linkage disequilibrium, the nonrandom association between all the variants in the block.
- Common disease susceptibility alleles are of ancient origin and are mildly deleterious (such as regulatory sequence mutations and weak missense mutations). They avoid being eliminated by purifying selection by having little effect on reproductive rates, or by simultaneously conferring some selective advantage now, or in the past.
- GWA studies have successfully identified thousands of disease-associated SNP markers; because almost all are of weak effect, they have limited use in predicting disease risk.
- Copy number variants (CNVs) do not generally make a large contribution to genetic susceptibility to disease but do occur at increased frequencies in some disorders.
- GWA studies have been of great value in elucidat ing the biological pathways in complex diseases, with prospects for identifying new drug targets and treatments.
- Nongenetic factors are clearly very important in com plex diseases, but standard case-control studies are limited in their ability to detect gene–environment interactions. Prospective cohort studies are more suited to that task; they study individuals over a long timeframe that commences before the onset of disease.
- Environmental factors work at different levels to influ ence disease susceptibility. One important way is to alter the epigenetic settings of cells, resulting in altered gene expression. Altered epigenetic settings in early life are

thought to alter the risk of various adult diseases such as diabetes and cardiovascular diseases.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

General genetic mapping and meiotic recombination frequency

Altshuler D, Daly MJ & Lander ES (2008) Genetic mapping in human disease. *Science* 322:881–888; PMID 18988837.

Cheung VG (2007) Polymorphic variation in human meiotic recombination. *Am J Hum Genet* 80:526–530; PMID 17273974.

Coop G (2008) High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science* 319:1395–1398; PMID 18239090.

Ott J (1999) *Analysis of Human Genetic Linkage*, 3rd ed. Johns Hopkins University Press. [Authoritative, detailed account.]

Gene identification in Mendelian disorders

Bamshad MJ (2011) Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 12:745–755; PMID 21946919.

Puliti A (2007) Teaching molecular genetics: chapter 4—positional cloning of genetic disorders. *Pediatr Nephrol* 22:2023–2029; PMID 17661092.

Heritability and heritability studies

Lichtenstein P (2009) Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 373:234–239; PMID 19150704.

Visscher PM (2008) Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet* 9:255–266; PMID 18319743.

Wells JCK & Stock JT (2011) Re-examining heritability: genetics, life-history and plasticity. *Trends Genet* 10:421–428; PMID 21757369.

Quantitative traits and the liability/threshold model

Falconer DS (1965) The inheritance of liability to certain diseases estimated from the incidence among relatives. *Ann Hum Genet* 29:51–76; doi 10.1111/j.1469–1809.1965.tb00500.x. [The original formulation of the liability threshold model to explain dichotomous traits.]

Lango Allen H (2010) Hundreds of variants clustered at genomic loci and biological pathways affect human height. *Nature* 467:832–838; PMID 20881960.

Linkage analysis in complex disease

Altmüller J (2001) Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 69:936–950; PMID 11565063.

Hugot JP (1996) Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 379:821–823; PMID 8587604.

Risch N & Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273:1516–1517; PMID 8801636.

Weeks DE (2004) Age-related maculopathy: a genomewide scan with continued evidence of susceptibility loci within the 1q31, 10q26, and 17q25 regions. *Am J Hum Genet* 75:174–189; PMID 15168325.

Linkage disequilibrium, haplotype blocks, and the HapMap Project

Ardlie KG (2002) Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 3:299–309; PMID 11967554.

Daly MJ (2001) High-resolution haplotype structure in the human genome. *Nat Genet* 29:229–232; PMID 11586305.

The International HapMap Consortium (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449:851–862; PMID 17943122.

Slatkin M (2008) Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* 9:477–485; PMID 18427557.

General reviews on genetic association and GWA studies

Lewis CM & Knight J (2012) Introduction to genetic association studies. *Cold Spring Harbor Protoc* 3:297–306; PMID 22383645.
Manolio TA (2013) Bringing genomewide association findings into clinical use. *Nat Rev Genet* 14:549–558; PMID 23835440.
Tam V (2019) Benefits and limitations of genomewide association studies. *Nat Rev Genet* 20:467–484; PMID 31068683.
Visscher PM (2017) 10 years of GWAS discovery: biology, function and translation. *Am J Hum Genet* 101:5–22; PMID 28686856.

Development and technical aspects of GWA studies

Browning SR & Browning BL (2011) Haplotype phasing: existing methods and new developments. *Nat Rev Genet* 12:703–714; PMID 21921926.
Caliskan M (2021) A catalog of GWAS fine-mapping efforts in autoimmune disease. *Am J Hum Genet* 108:549–563; PMID 33798443.
Cooper GM & Shendure J (2011) Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat Rev Genet* 12:628–640; PMID 21850043.
GWAS Catalog. *The NHGRI-EBI Catalog of human genomewide association studies*. Available at www.ebi.ac.uk/gwas/ (a general background description is also available at PMID 30445434).
Marigorta UM (2018) Replicability and Prediction: Lessons and Challenges from GWAS. *Trends Genet* 34:504–517; PMID 29716745.
Wellcome Trust Case Control Consortium (2007) Genomewide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* 447:661–678; PMID 17554300.

Missing heritability, evolution of common variants, and rare variants and copy number variants in complex disease

Manolio TA (2009) Finding the missing heritability of complex diseases. *Nature* 461:747–753; PMID 19812666.
Nakagome S (2012) Crohn's disease risk alleles on the NOD2 locus have been maintained by natural selection on standing variation. *Mol Biol Evol* 29:1569–1585; PMID 22319155.
Raychaudhuri S (2011) Mapping rare and common causal alleles for complex human diseases. *Cell* 147:57–69; PMID 21962507.
Girirajan S (2011) Human copy number variation and complex disease. *Annu Rev Genet* 45:203–226; PMID 21854229.

Malhotra D & Sebat J (2012) CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell* 148:1223–1241; PMID 22424231.

Singh T (2022) Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* 604:509–516; PMID 35396579.

Wainschtein P (2022) Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. *Nature Genet* 54:263–273; PMID 35256806.

Genetic risk prediction in complex disease

Choi SW (2020) Tutorial: a guide to performing polygenic risk score analyses. *Nat Protocol* 15:2759–2772; PMID 32709988.

Khera AV (2018) Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 50:1219–1224; PMID 30104762.

Visscher PM & Gibson G (2013) What if we had whole-genome sequence data for millions of individuals? *Genome Med* 5:80; PMID 24050736.

Visscher PM (2021) Discovery and implications of polygenicity of common disease. *Science* 373:1468–1473; PMID 34554790.

Wray NR (2010) The genetic interpretation of area under the ROC curve in genomic profiling. *PLOS Genet* 6:e1000864; PMID 20195508.

Genetic architecture and biological pathways in complex disease

Bertram L & Tanzi RE (2012) The genetics of Alzheimer's disease. *Prog Mol Biol Transl Sci* 107:79–100; PMID 22482448.

Khor B, Gardet A & Xavier RJ (2011) Genetics and pathogenesis of inflammatory bowel disease. *Nature* 474:307–317; PMID 21677747.

King RA, Rotter JI & Motulsky AG (eds) (2002) *The Genetic Basis of Common Disease*, 2nd ed. Oxford University Press.

Sullivan PF (2012) Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet* 13:537–551; PMID 22777127.

Epigenetic effects and gene-environment interactions in complex disease

Cavalli G & Heard E (2019) Advances in epigenetics links genetics to the environment and disease. *Nature* 571:489–499; PMID 31341302.

Czyz W (2012) Genetic, environmental and stochastic factors in monozygotic twin discordance with a focus on epigenetic differences. *BMC Med* 10:93; PMID 22898292.

Feil R & Fraga MF (2012) Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet* 13:97–109; PMID 22215131.

Grossniklaus U (2013) Transgenerational epigenetic inheritance: how important is it? *Nat Rev Genet* 14:228–235; PMID 23416892.

Kaelin WG Jr & McKnight SL (2013). Influence of metabolism on epigenetics and disease. *Cell* 153:56–69; PMID 23540690.

Manolio TA (2006) Genes, environment and the value of prospective cohort studies. *Nat Rev Genet* 7:812–820; PMID 16983377.

Portela A & Esteller M (2010) Epigenetic modifications and human disease. *Nat Biotechnol* 28:1057–1068; PMID 20944598.

Virgin HW & Todd JA (2011) Metagenomics and personalized medicine. *Cell* 147:44–56; PMID 21962506. [Considers how the human microbiome influences disease risk.]

9

Genetic approaches to treating disease

DOI: 10.1201/9781003044406-9

CONTENTS

Treatment of genetic disease and genetic treatment of disease are two separate matters. The cause of a disease (whether mostly genetic or mostly environmental) and its treatability are quite unconnected. Standard medical treatments to alleviate disease symptoms—hearing aids or cochlear implants for treating profound deafness, for example—are just as applicable if the disease is mostly genetic or mostly environmental. In this chapter the primary focus is on how genetic technologies are being applied to treat disease, but we begin by taking a broader look at different treatment strategies for genetic disorders.

For the great majority of genetic conditions, even for single-gene disorders, existing treatments are lacking or unsatisfactory. **Figure 9.1** represents a snapshot taken in 1999 when the treatability of 372 genetic diseases was assessed by Charles Scriver and Eileen Treacy. The situation has improved since then, but we still have a long way to go.

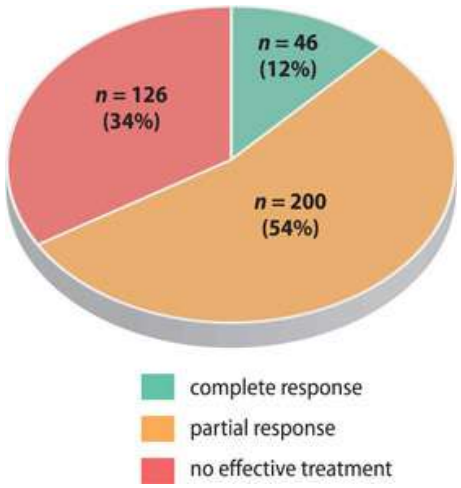


Figure 9.1 Treatment of genetic disorders has often not been very effective. The data here record the response to treatment for 372 single-gene disorders in which the underlying gene or its biochemical function were known and representative treatment data were available by 1999. (Data from Scriver CR & Treacy EP [1999] *Mol Genet Metab* 68:93–102; PMID 10527662.)

Causative genes for many single-gene disorders have been identified comparatively recently, and it may take many years of research to identify how the underlying genes function normally in cells and tissues. Armed with that knowledge, we might hope to devise better treatments in the future, including those that target the *cause* of disease, rather than just dealing with the symptoms. Towards that end, a variety of biological drugs/treatments (sometimes collectively known as *biologics*) have been devised including the use of therapeutic proteins and RNAs, plus gene therapies that offer the opportunity for highly effective treatment for certain monogenic disorders; for other monogenic disorders there may be difficult technical obstacles to devising effective biological therapies—we provide examples below. But despite their promise, biological drugs/treatments can be very expensive, and as genetic causes of disease are understood more precisely new conventional drug therapies may be devised—we give examples below.

We cover treatments for cancers in Chapter 10. For other complex diseases, reasonably satisfactory treatments may exist, such as in the case of diabetes; for many others the treatments are less than satisfactory, or ineffective. By definition, complex diseases are complex at the genetic level: only a decade ago we knew very few of the underlying genetic factors, but recent studies have since revealed many of the contributing genetic factors. In some cases, genetic studies will be able to divide individual complex diseases into subtypes (*disease stratification*), allowing different treatments to be tailored to suit different disease subtypes. The emerging information will place us in a better position to develop novel, more effective treatments.

Environmental factors are clearly very important in complex diseases and have been notably well documented in many cancers. Some environmental factors are also well recognized in some noncancer conditions. Cigarette smoking is a powerful factor in age-

related macular degeneration and emphysema, for example, and the importance of a healthy diet and regular exercise is well recognized in conditions such as type 2 diabetes. Considerable work needs to be done to extend our knowledge of contributory environmental factors. That will provide opportunities for effective interventions, because exposure to an environmental factor can often be modified.

In this chapter we are primarily concerned with molecular approaches to treating disease. In Section 9.1 we give an overview. First, we look at how treatments can be classified into different categories. We take a broad view of the different levels at which disease can be treated, and explore the different genetic technology inputs that can be applied. In Section 9.2 we cover genetic inputs to treating disease with synthetic hydrocarbon-based chemical drugs ("small molecule drugs") and therapeutic proteins, including genetically altered antibodies, and we explore aspects of pharmacogenetics that deal with the different responses of patients to small molecule drugs and aspects of drug metabolism. Variation in how we respond to chemical drugs is very important: it leads to hundreds of thousands of fatalities per year.

In Section 9.3 we cover the principles and general methodology of different therapeutic methods involving the genetic modification of a patient's cells (gene therapy). In this section we also describe related stem cell therapy methods. All the methods described above need to be tested in animal disease models before clinical trials are carried out, and we briefly deal with different approaches to disease modeling in the closing section of Section 9.3.

Finally, in Section 9.4 we describe how gene therapy has been applied in clinical trials, assess the progress made, and consider future prospects, including the use of therapeutic RNAs. Ethical issues are considered later, in Chapter 11.

9.1 AN OVERVIEW OF TREATING GENETIC DISEASE AND OF GENETIC TREATMENT OF DISEASE

In this introductory section we first look at broad categories of treating genetic disease and illustrate the diversity of treatments with examples from inborn errors of metabolism. We then consider the different levels at which molecular-based disease treatments can be applied.

Three different broad approaches to treating genetic disorders

Two types of treatment can be used, according to whether pathogenesis is due to some defined genetic deficiency or some positively harmful effect (rather than a *lack* of some important gene product or metabolite). A third type of treatment seeks to reduce susceptibility to disease by understanding the pathway involved (**Figure 9.2**). We expand on these themes in the sections below, taking into account both current practice and experimental therapies.

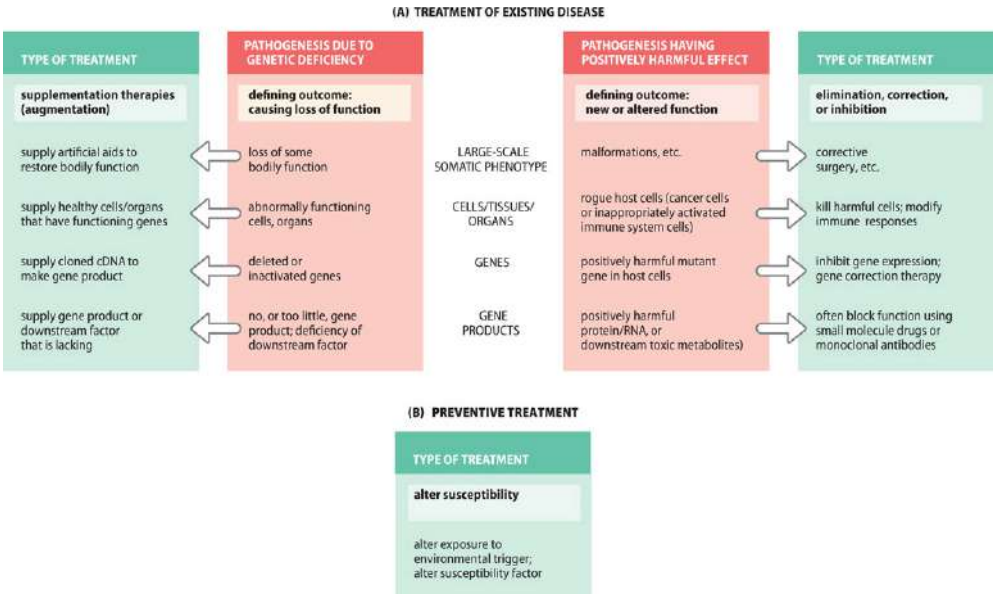


Figure 9.2 Different major treatment strategies for genetic disorders. Note that some strategies are experimental. (A) Different types of treatment for an existing genetic condition according to whether the disease is caused by a genetic deficiency or some other genetic effect, and according to the level at which treatment is applied. Supplementation (augmentation) therapies seek to compensate for a genetic deficiency in various ways: by supplying purified functional gene product directly (protein supplementation); by supplying a purified downstream factor that is lacking; or by indirectly supplying either cloned DNA or healthy cells (either from a donor, or genetically modified cells from the patient) to make the missing gene product. Therapies for conditions where the pathogenesis has a positively harmful effect, may work at different levels. At the cell/tissue level, the object is to deal with rogue cells that behave abnormally to cause disease (notably cancer cells, or immune system cells that attack host cells in autoimmune and inflammatory diseases). At the gene level, the object is to prevent the harmful effects of a gain-of-function mutation (or a gene from a pathogenic microorganism), and at the gene product level, the aims may be to eliminate or reduce the production of elevated toxic metabolites, as in some inborn errors of metabolism. (B) Disease prevention strategies include altering exposure to environmental triggers, such as through extreme dietary modifications in some inborn errors of metabolism, and the use of drugs such as statins to alter disease susceptibility.

Supplementation therapy for genetic deficiencies

In some genetic disorders the problem is the loss of some normal function. In principle, these disorders might be treated by *supplementation (augmentation) therapy*: something is provided to the patient that *supplements* a severely depleted, or missing, factor, thereby overcoming the deficiency and restoring function. Different types of supplement can be provided to restore function at different levels. At the level of the somatic phenotype, treatment can be conventional—providing cochlear implants or hearing aids to treat hereditary deafness, for example.

At the molecular level, the phenotype can be restored by providing a purified gene product that is lacking—a missing enzyme, say, in many inborn errors of metabolism. Alternatively, when the gene product works in a biological pathway required to synthesize some important downstream factor, such as a lipid hormone, it might be a lack of the downstream factor that is treated (by providing purified lipid hormone, in this case). At a higher molecular level, the aim of some types of supplementation (augmentation) therapy has been to transfer a cloned cDNA into the tissues of a patient where it can be expressed to make a missing protein.

At the cellular and organ levels, healthy cells and organs can be transplanted into a patient to make a product that the patient lacks, or to compensate for deficiency of a particular cell type. That can involve transplanting cells from a donor, as in bone marrow transplantation or organ transplantation. More recently, some cellular gene therapies have been used very successfully; here, the cells of the patient are genetically modified so that they can now express the desired gene product. Novel stem cell therapies seek to treat disease by supplying cells of a particular type that are lacking.

Applicability of molecular supplementation therapy

Recessive disorders (where both alleles have lost their function) are more suited to molecular supplementation therapy than are dominant disorders. Affected individuals quite often cannot make any functional copies of some normal gene product. Even a modest efficiency in delivery (of healthy cells, genes, or proteins) to an affected individual can often allow effective treatment, and recently there have been substantial breakthroughs. As illustrated below, however, supplementation gene therapy is currently not practical for some recessive disorders—it can often be difficult to get efficient delivery and production of the desired molecules.

In dominant disorders due to haploinsufficiency, disease occurs even when one allele is normal and present in all diploid cells; efficient delivery and production of the missing gene product would be essential. In addition, the underlying gene is dosage sensitive (described in Box 7.3). Very precise supplementation therapy—a very difficult prospect—would be needed, and is currently unavailable. Supplementation therapy can also be applied to certain complex diseases, for example by treating diabetes using purified insulin, or by transplantation of pancreatic islet cells.

Treatment for disorders producing positively harmful effects

A second, different approach to treatment is needed for diseases in which the pathogenesis involves a positively harmful effect, rather than a deficiency. Here, supplementation therapy cannot be used: something has gone wrong that cannot be corrected by simply administering

some normal gene, normal gene product, or normal cells to the patient. Different methods are needed (see Figure 9.2B).

The harmful effect might be treatable at the somatic phenotype level, as in the case of some developmental malformations: corrective surgery is highly effective, for example, in treating various complex disorders such as congenital heart defects, cleft lip and palate, and pyloric stenosis.

At the molecular level, treatments can be conducted at different stages. In many inborn errors of metabolism the problem is elevated levels of harmful metabolites that can be tackled in different ways, as described in the next major section. A more general problem is presented by actively harmful gene products from a mutant gene. Examples include mutant prion proteins and b-amyloid, which are liable to form protein aggregates that can kill cells (described in Clinical Box 8 on page 229), and also harmful proteins or RNAs formed after the unstable expansion of short oligonucleotide repeats (detailed in Section 7.2). Dangerous mutant gene products may be combatted by using a small molecule drug or therapeutic monoclonal antibody to bind selectively to the mutant molecule and inhibit its activity.

In some cases, the therapeutic strategies are used to selectively inhibit the expression of a harmful gene at the mRNA level, as described in detail in Section 9.3. In addition, at the gene level, experimental corrective gene therapy has the potential to repair damaged genes, replacing a pathogenic mutation by a normal sequence, as described below.

At the cellular level, the problem may manifest itself as harmful cells. Some mutations can induce cells to behave abnormally, proliferating excessively to cause cancers that have been treated by long-standing methods (surgical excision, radiation, and chemotherapy) and more recently by targeted chemical and biological drugs (we cover these plus the prospects of cancer gene therapies in Chapter 10). In some genetic disorders, the problem is excessive immune responses in which certain immune system cells inappropriately attack host cells (in autoimmune disorders such as rheumatoid arthritis, and in inflammatory diseases such as Crohn's disease). Here, there is the potential to employ therapies that down-regulate immune responses, but in some cancer gene therapy trials the exact opposite approach has been taken (upregulating immune responses in an attempt to kill cancer cells).

Treatment by altering disease susceptibility

A third way of treating disease seeks to reduce susceptibility to disease in some way, for certain monogenic disorders and also some complex diseases. In some inborn errors of metabolism, blockage of one step in a metabolic pathway can drive alternative pathways, causing a buildup of toxic metabolites. A solution here may be to reduce disease susceptibility by removing an environmental trigger. And in some diseases, key susceptibility factors can be manipulated to reduce the chance of disease recurrence, or the effects of a progressive disease.

Very different treatment options for different inborn errors of metabolism

Inborn errors of metabolism, founded on the work of Archibald Garrod on alkaptonuria in the early 1900s, were the first genetic disorders to be understood at the biochemical level. Since then, we have developed a detailed understanding of the molecular pathology for many of these disorders.

Early optimism that disease treatments would follow once we knew all the major details of pathogenesis has suffered quite a setback: effective treatment is unavailable for many of these disorders. But there has been a steady improvement. In a longitudinal study of 65 inborn errors of metabolism published in 2008, significant improvement was recorded: 31 disorders showed no signifi-cant response to treatment in 1983 but only 17 in 2008; and the number of conditions fully responding to treatment jumped from 8 in 1993 to 20 in 2008. As we describe later, there have been some important and encouraging successes using various therapies in the last dozen years.

In this section, we use inborn errors of metabolism to illustrate the different ways in which treatment can be offered (each of the three general strategies shown in Figure 9.2 has been successfully applied), and why effective treatment will be difficult for some disorders.

Two broad phenotype classes

For the simplest cases, picture a metabolic pathway composed of sequential steps, each catalyzed by an individual enzyme. Now imagine that loss-of-function mutations have inactivated the gene encoding one of the enzymes. The resulting absence of that enzyme will lead to a lack of downstream product plus a buildup of substrate proximal to (before) the blocked step.

Sometimes, and often in biosynthetic pathways, the most noticeable effect is the lack of end product. In these cases, supplementation therapy can compensate for a lack of a gene product, or of some other downstream molecule whose production depends on the gene product.

In other cases, as previously described for phenylketonuria in Clinical Box 9 on page 234, the buildup of precursors proximal to the blocked step drives alternative pathways, producing abnormal concentrations of some metabolites that have harmful effects. While a disorder such as this is caused by a recessive loss of function at the disease locus, the disease phenotype results from what is at the cellular and physiological levels, a gain of function: a buildup of positively harmful metabolites that requires different treatment strategies. As we show below, the disease phenotype for some disorders may have components that are treatable by supplementation therapy, and others that are due to positively harmful effects.

Supplementation therapy

Here the missing product is provided to overcome the deficiency. It might be the gene product itself (protein supplementation) or a critically important downstream factor that it regulates,

which may not be a protein. For example, recessive congenital hypothyroidism (OMIM 275200) is due to a deficiency in thyroid hormone (a *lipid* hormone whose production is largely dependent on a peptide hormone, thyroid-stimulating hormone). Affected individuals have a mutant thyroid-stimulating hormone receptor that fails to respond to thyroid-stimulating hormone, but they are effectively treated with purified thyroid hormone. Mutations in *CYP21A2* cause 21-hydroxylase deficiency by disturbing the production of steroid hormones. The effects include greatly reduced levels of two types of lipid hormones: steroid hormones of the mineralocorticoid class (such as aldosterone) and glucocorticoid class (such as cortisol); the deficiencies in these hormones can be treated by relevant steroid supplementation therapies (**Figure 9.3A**).

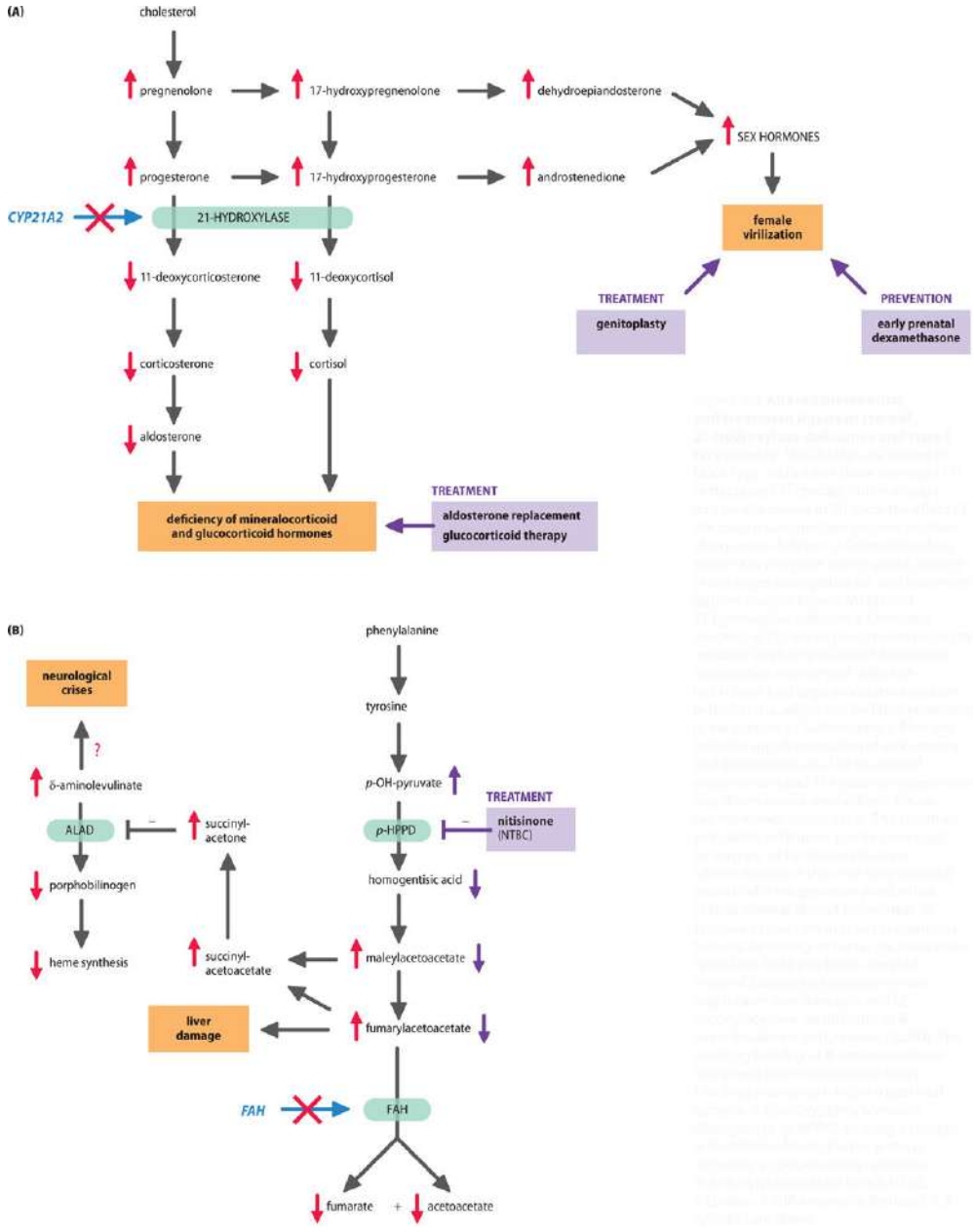


Figure 9.3 Altered metabolism and treatment inputs in steroid 21-hydroxylase deficiency and type 1 tyrosinemia. Metabolites are shown in black type; red arrows show increased (↑) or decreased (¯) production in disease and purple arrows in (B) show the effect of the drug nitisinone. Red crosses: location of a genetic deficiency. Colored shading shows: key enzymes (green ovals), disease phenotypes (orange boxes), and treatment options (purple boxes). (A) Steroid 21-hydroxylase deficiency. Complete absence of 21-hydroxylase results in greatly reduced levels of the steroid hormones, aldosterone and cortisol. Affected individuals lose large amounts of sodium in their urine, which can be life-threatening in early infancy ("salt-wasting"). Therapy includes supplementation of aldosterone and glucocorticoids. The buildup of progesterone and 17-hydroxyprogesterone also drives excess production of male sex hormones (androgens). The resulting virilization in females can be corrected by surgery, or by dexamethasone administration in the very early prenatal period (which suppresses production of fetal adrenal steroid hormones). (B) Tyrosine catabolism in type I tyrosinemia. Genetic deficiency of fumarylacetoacetate hydrolase (FAH) produces elevated levels of fumarylacetoacetate (which may induce liver damage), and of succinylacetone, an inhibitor of ***δ***-aminolevulinate dehydratase (ALAD). The resulting buildup of ***δ***-aminolevulinate may precipitate neurological crises. The drug nitisinone inhibits a proximal enzyme, *p*-hydroxyphenylpyruvate dioxygenase (*p*-HPPD), causing a change in metabolite levels (purple arrows), including a compensatory reduction in fumarylacetoacetate levels. NTBC, 2-(2-nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione.

Supplementation therapy can also involve transplanting cells or organs to supplement a genetic deficiency. Bone marrow transplantation, effectively a way of transplanting hematopoietic stem cells, has frequently been used to treat disorders of blood cells (or other cells originating from hematopoietic stem cells). Liver transplantation has been used for many serious inborn errors of metabolism (many metabolic enzymes are synthesized by the liver). Because of the possibility of graft rejection, organ transplantation is a serious matter; it is indicated when the disorder is expected to progress to organ failure—the treatment is intended to save lives, not just cure patients.

A more recent, and quite different, treatment uses a form of gene therapy: the gene product is obtained after the cells of a patient have been genetically modified to contain and express a functional copy of the relevant gene (as detailed in Sections 9.3 and 9.4). And stem cell therapy is designed to replace cells of a particular type that have been lost through disease (or injury).

Treating or preventing harmful effects of elevated metabolites

When abnormally elevated levels of metabolites cause disease, different treatments can be devised. One way is to use drugs to cause a compensatory change in metabolite levels. Take type 1 tyrosinemia (PMID 20301688). This disorder results from a deficiency of fumarylacetoacetate hydrolase, which catalyzes the terminal step in the tyrosine degradative pathway (Figure 9.3B). The resulting buildup of precursors leads to liver and renal tubule dysfunction, and untreated children may have repeated neurologic crises. Oral administration

of nitisi-none (also called NTBC) provides effective treatment—it inhibits a proximal (upstream) enzyme step, causing a compensatory reduction in fumarylacetoacetate hydrolase levels (Figure 9.3B).

Normal levels of elevated metabolites can be restored by the removal of excess amounts from the body. Phlebotomy is a possibility if the excess metabolite is present in blood (**Table 9.1**). Alternatively, some indirect means can be used to force the metabolite into another metabolic pathway to decrease the levels to normal values—see the example of dealing with excess ammonia in urea cycle disorders in **Figure 9.4**.

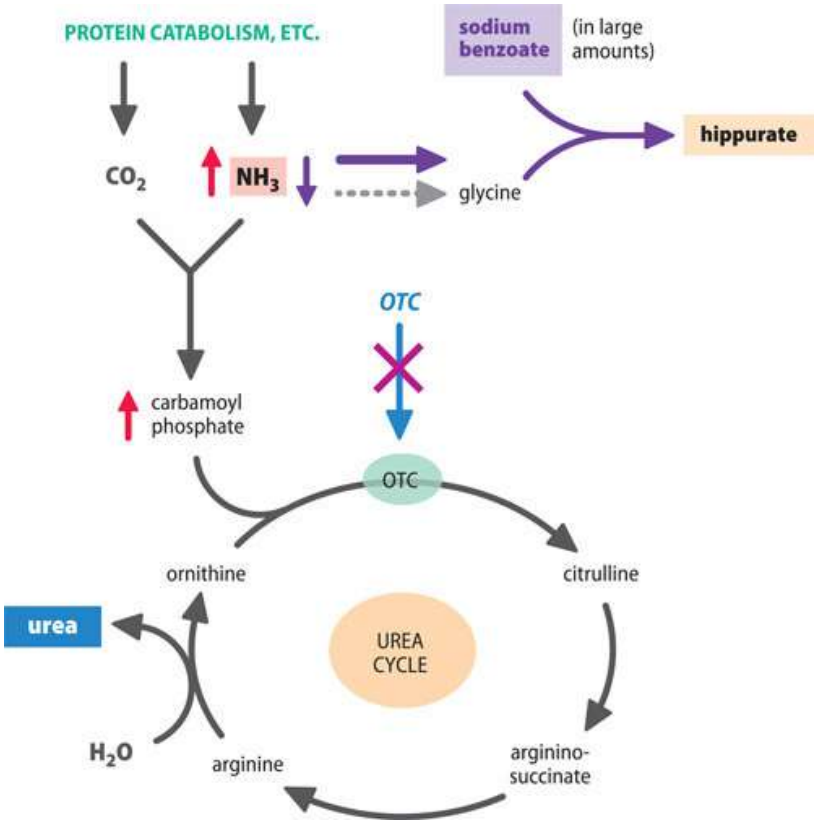


Figure 9.4 Reducing elevated metabolite levels by shunting the metabolite into an alternative metabolic pathway. The urea cycle normally serves to convert ammonia (NH3), which is neurotoxic, to nontoxic urea. But in urea cycle disorders, ammonia cannot be converted to urea and builds up. In ornithine transcarbamylase (OTC) deficiency (OMIM 311250), the metabolic block causes an increase in levels of the proximal metabolites, carbamoyl phosphate and ammonia (vertical red arrows). The therapy here involves treating a patient with large amounts of sodium benzoate and takes advantage of a normally minor pathway in which some ammonia is naturally converted into small amounts of glycine. Benzoate ions conjugate with glycine to form hippurate, which is excreted in urine. By removing glycine, the treatment drives the production of replacement glycine from ammonia (thick purple horizontal arrow), thereby reducing ammonia levels (vertical purple arrow).

TABLE 9.1 EXAMPLES OF DIFFERENT TYPES OF TREATMENT OF INBORN ERRORS OF METABOLISM

Treatment	Type of action	Examples and comments

Treatment	Type of action	Examples and comments
Supplementation (augmentation) therapy	protein supplementation	enzyme replacement therapies for many inborn errors of metabolism; provision of blood clotting factors in hemophilias (see Section 9.2 for details)
	hormone replacement	thyroid hormone for infants with congenital hypothyroidism; growth hormone for growth hormone deficiency (see also Figure 9.3A)
	bone marrow transplantation	useful for disorders affecting blood cells and some other immune system cells (as illustrated in Figure 9.20), such as mucopolysaccharidosis type 1 (Hurler syndrome, OMIM 607014)
	organ transplantation	liver transplantation has been used successfully for various inborn errors of metabolism, including α_1-antitrypsin deficiency and urea cycle disorders
	gene supplementation	successfully used for different types of severe combined immunodeficiency, hemophilia (Sections 9.3 and 9.4)
Counteracting harmful effects of abnormally elevated metabolites	manipulated excretion of metabolite	periodic phlebotomy (blood removal) is a very effective treatment for directly removing excess iron in the iron overload condition hemochromatosis (OMIM 235200)
	shunting of elevated metabolite into side metabolic pathway	for urea cycle disorders the buildup of toxic ammonia can be alleviated by sodium benzoate treatment, driving excess ammonia to be metabolized in a side pathway (Figure 9.4)
	inhibition of a proximal step in a pathway leading to harmful metabolites	babies with type 1 tyrosinemia (OMIM 276700) are unable to metabolize tyrosine effectively and suffer liver damage from toxic intermediates. The drug NTBC inhibits a proximal enzyme to prevent buildup of toxic intermediates (Figure 9.3B)
Prevention (avoiding or reducing susceptibility)	substrate restriction (diet modified to severely reduce or eliminate intake of substrate for a deficient enzyme)	reduced intake of phenylanine in phenylketonuria. Elimination of galactose in galactosemia (OMIM 230400); affected individuals completely lack galactose-1 -phosphate uridyltransferase. Galactose is a component of the lactose in milk but is inessential, and milk is completely withdrawn from the diet

Treatment	Type of action	Examples and comments
	reduction of a susceptibility factor	in familial hypercholesterolemia, *LDLR* gene mutations result in low levels of the low-density lipoprotein receptor; the resulting elevated plasma low-density lipoprotein cholesterol levels predispose to cardiovascular disease but can be effectively lowered by statins, drugs that inhibit a proximal enzyme, HMCoA reductase, in the cholesterol biosynthesis pathway

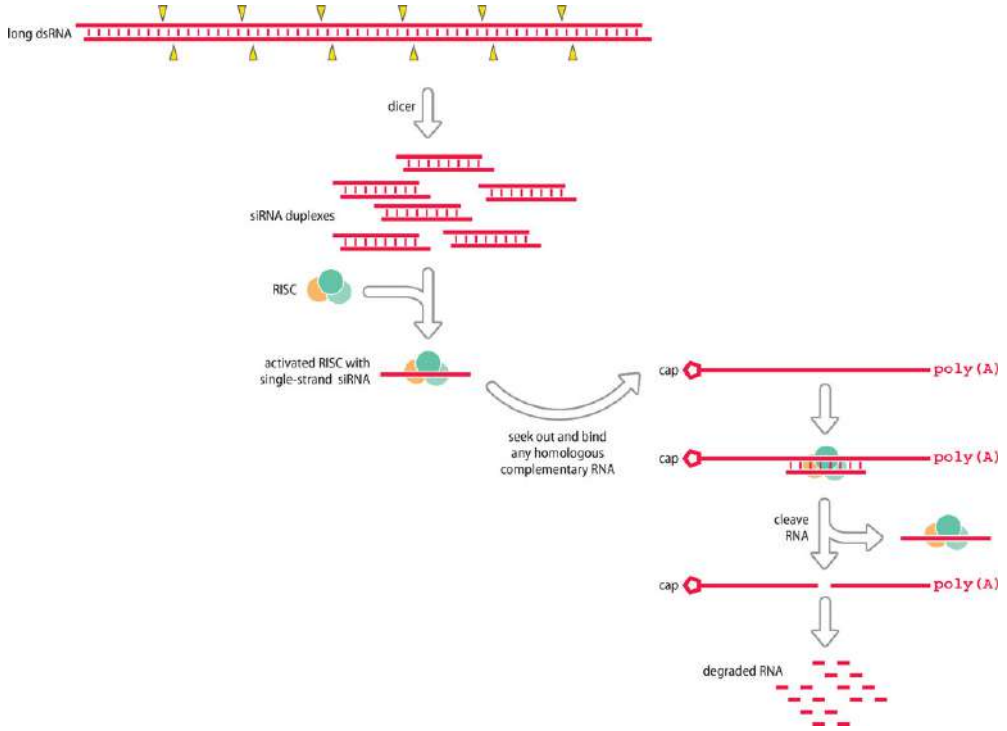


Figure 6.13 RNA interference (RNAi). Long double-stranded (ds) RNA is an unusual structure in our cells, but may signal the presence of some invading viruses or excess transposon activity. To defend cells against these threats, Dicer, a cytoplasmic ribonuclease, cleaves dsRNA asymmetrically on the two strands at positions 21 nucleotides apart (yellow triangles). The resulting short interfering RNA (siRNA) consists of a duplex of two 21-nucleotide sequences with overhangs of two nucleotides at the 3¢ ends. The siRNA duplexes are bound by RNA-induced silencing complexes (RISC) that degrade one of the siRNA strands to give an activated RISC complex with a single siRNA strand (called the *guide* RNA). The RISC-siRNA complex is now activated and will bind (by RNA–RNA base pairing) to any RNA sequence that is complementary in sequence to the guide RNA (such as a specific viral mRNA sequence). The cleaved RNA fragments, lacking a protective cap or poly(A) sequence, are vulnerable to attack by cellular exonucleases and are rapidly degraded. Note that introducing long dsRNA into mammalian cell culture results in the indiscriminate destruction of mRNA, so siRNAs need to be short, often 21 nucleotides long.

A different approach, *disease prevention*, seeks to *prevent* the buildup of toxic levels of metabolite. In some cases, *substrate restriction* is used: the diet is modified to severely reduce or eliminate the intake of a substrate of the deficient enzyme. That can work very successfully when the blocked enzyme is at the start of a pathway that metabolizes a dietary component. This approach prevents the harmful effects of toxic metabolites that build up in phenylketonuria (as described in Clinical Box 9 on page 234). Even in this case, the treatment requires lifelong compliance with a rather restricted, and difficult, diet. Prevention is sometimes possible at the prenatal level, as in the case of treating virilized female fetuses in 21-hydroxylase deficiency (Figure 9.3A).

Mixed success in treatment

Treatment of some inborn errors of metabolism is very effective, as with phenylketonuria. However, the treatment can be sub-optimal in some cases, and very difficult or essentially nonexistent in others for various reasons. If a disorder is congenital and harmful effects have occurred during development, treatment options may be limited. In some cases, potential therapy can be frustrated by delivery problems. In Tay–Sachs syndrome (PMID 20301397), for example, deficiency in hexosaminidase A leads to an inexorable buildup of a sphingolipid GM2 ganglioside to toxic levels, causing damage to brain cells.

Genetic treatment of disease may be conducted at many different levels

Any disease, whether it has a genetic cause or not, is potentially treatable using a range of different procedures that apply genetic manipulations or genetic knowledge in some way (**Figure 9.5**).

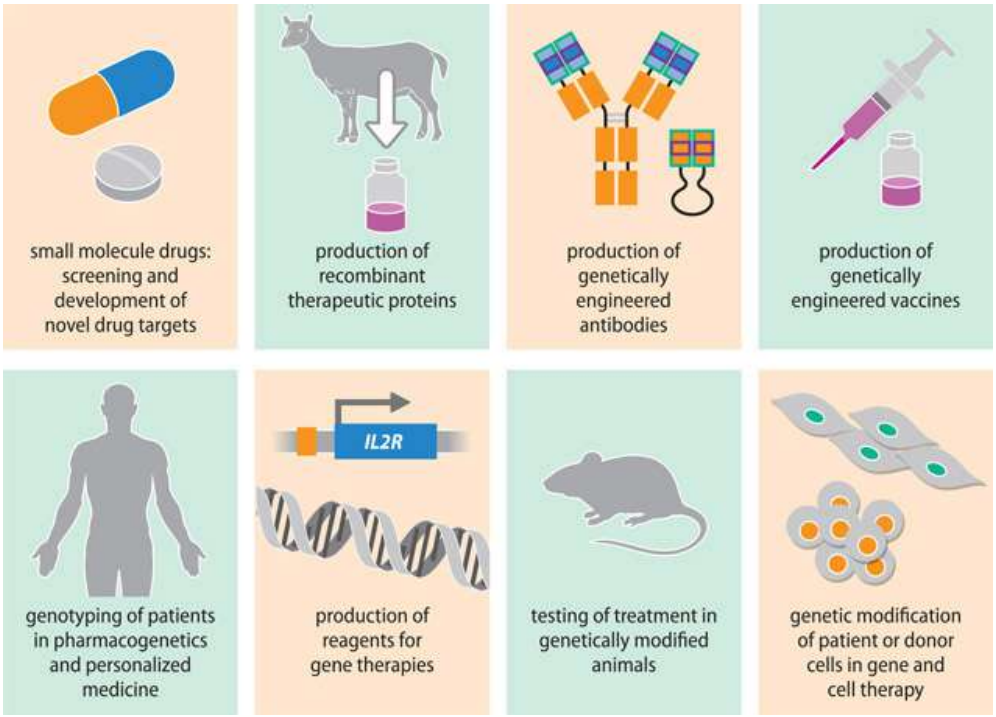


Figure 9.5 Some of the many different ways in which genetic technologies are used in the treatment of disease.

Sometimes genetic techniques form part of a treatment regime that also involves conventional small molecule drugs or vaccines.

Pharmacogenetics is concerned with how the actions of drugs and the reactions to them vary according to variation in the patient's genes. Genotyping of individuals might then be used to predict patterns of favorable and adverse responses to specific drug treatments. Such genotyping may become routine as massively parallel DNA sequencing permits extensive screening of genes in vast numbers of people.

New targets for drug development are being identified using a knowledge of genetics and cell biology. Genetic techniques can also be used directly in producing drugs and vaccines for treating disease. Another active area concerns treating disease with therapeutic proteins that are produced or modified by genetic engineering. Genes are cloned and expressed in suitable cultured cells or organisms to make large amounts of a specific protein that is then purified *(recombinant proteins),* including hormones, blood factors, and enzymes, and especially genetically engineered antibodies.

Gene therapies are the ultimate genetic application in treating disease; they rely on genetically modifying the cells of a patient. Delivering therapeutic constructs into the stem cells of a patient is particularly valuable when the disease primarily affects short-lived cells, such as blood cells. Animal models are especially important resources for testing new therapies before they are used in clinical trials. As described in Section 9.3, the vast majority of animal models of disease have been generated by the genetic manipulation of rodents, notably mice.

9.2 GENETIC INPUTS INTO TREATING DISEASE WITH SMALL MOLECULE DRUGS AND THERAPEUTIC PROTEINS

Chemical treatments for disease are developed by the pharmaceutical industry. Previously they relied almost exclusively on hydrocarbon-based small molecule drugs synthesized by standard chemical reactions. More recently they have been joined by a new class, biological drugs (*biologics*), including therapeutic proteins that have been prepared by genetic engineering. We give a brief description of the two classes below. Then we cover the important question of how genetic variation between people can result in very wide differences in how we respond to therapeutic drugs, and consider aspects of drug metabolism. We finish with examples of therapeutic uses of the two principal drug classes.

Small molecule drugs

The conventional drug discovery process has involved screening huge numbers of small synthetic molecules with hydrocarbon backbones for evidence that they can reduce pathogenic effects. A drug such as this typically works by binding to a specific target protein with a key role in the pathogenesis—often a receptor, ion channel, or enzyme. The drug is able to bind to the protein by fitting into some cleft, groove, or pocket at a key position in the protein structure. Binding of the drug to the target protein may often prevent the protein from interacting with other molecules (and so block its function), or it may change the function of the protein in some way.

The drug screening process normally begins with preclinical assays in cell culture and in animal models to see whether a candidate drug has some encouraging properties. Promising candidates may be used in clinical trials, when their potential usefulness is monitored in different ways (**Figure 9.6**). To bring a drug to market is both costly (about US $1 billion) and time-consuming (often 12 years or more). Sometimes, however, a drug previously developed to treat one type of disease can be used to treat other diseases (*drug repurposing*); that is valuable because the drug has already been through lengthy and expensive clinical trials to assess its safety profile.

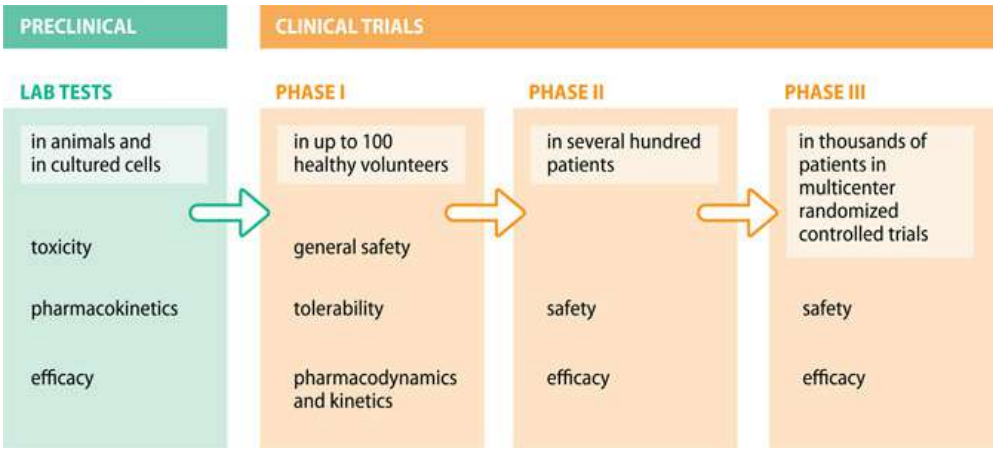


Figure 9.6 Major stages in drug development. The lists indicate the principal parameters tested at each stage. Pharmacokinetic testing assesses the absorption, activation, metabolism, and excretion of drugs. Pharmacodynamics monitors what the drug does to the body. Successive stages toward regulatory approval are increasingly expensive. To avoid unnecessary expenditure, any effects of genetic variation among patients that might influence marketability need to be identified as early as possible in the process.

Small molecule drugs have been with us for some time. Two important questions are: how effective are they, and how safe are they? Although the therapeutic value of many small molecule drugs on the market is questionable, many others have undoubtedly been of great service. But individual drugs affect different people in different ways. As explained below, many of our drug-handling enzymes are polymorphic; genetic variation between individuals has an important influence on both the efficacy and the safety of drugs.

All drugs currently on the market act through only a few hundred target molecules (they were first developed when information about possible targets was scarce), and the declining number of new drug applications and approvals over the past few years has reflected a crisis in drug target identification and validation.

New approaches

Recently, both genomics and genetic engineering have been making an impact on drug development. Genomic advances offer a broader perspective on how genetic variation affects drug metabolism, and the ways in which people respond to drugs. They are also providing additional potential drug targets, as described below.

Genetic engineering has been applied to allow the production of large quantities of different biopharmaceuticals (or *biologics*), including therapeutic "recombinant" proteins and genetically engineered monoclonal antibodies. Many of these are currently licensed to treat various disorders.

An overview of how genetic differences affect the metabolism and performance of small molecule drugs

Small molecule drugs have been with us for some time, and accumulated data tell us how effective they are and how safe they are. Taking the first point, drugs vary widely in effectiveness. Even when a drug receives regulatory approval, it is rarely effective in 100 % of the patients to whom it is prescribed. Some people might need higher or lower doses to achieve the same therapeutic effect. For others, a drug might simply have no therapeutic effect at all.

Most of us have benefited from effective antibiotics and painkillers, and drugs such as statins and beta-blockers (for reducing the risk of heart disease) have been of very considerable value. Certain other drugs, especially those used to treat psychiatric disease, are much less effective. That can mean wasted time and money, and extra suffering for patients. Then there is the enormous problem concerning adverse reactions to drugs, as described below.

Some of these differential effects are due to environmental causes: a person's ability to absorb or metabolize a drug may be changed by illness or lifestyle. Sometimes adverse effects occur as a result of interactions between different combinations of drugs taken by a sick person. But many differences are due to genetic variation between people, and genetic variation in our drug-handling enzymes is often pronounced.

Natural selection fosters genetic variation in the genes encoding our drug-handling enzymes because a major role of these enzymes has been to deal with unusual exogenous chemicals *(xenobiotics)* in our diet and environment. Like immune system molecules that recognize foreign antigens, drug-handling enzymes need to deal with potentially dangerous invaders some of which are subject to genetic control (such as ingested plant or fungal metabolites that might be toxic). The genetic variation in drug-handling enzymes that affects how the body deals with drugs is simply a small part of how we respond to the extraordinary range of chemicals encountered in our diet and environment.

Pharmacogenetics is the study of the roles of specific genes in these effects (or *pharmacogenomics*, when taking a genome-wide perspective of the role of genetic variation). Genetic factors affect both drug metabolism and effect, as studied by:

- **Pharmacokinetics**—the study of what the body does with, and to, the drug (encompassing drug absorption, activation, metabolism, and excretion).
- **Pharmacodynamics**—the study of the target response: how the person is affected by the drug.

Different stages at which genetic variation influences drug metabolism

Drug transport and metabolism within the body involve a series of stages. The pharmacological effect of the drug will be measured by its effect on cells within the desired *target tissue* or organ, but drug delivery is mostly nonspecific. Usually cells throughout the body are exposed to the drug so that the drug can exert a beneficial effect on what often might

be a small group of desired target cells. The circulation is the delivery route. That might mean direct intravenous injection, but most drugs are given orally or through intramuscular injection. After oral ingestion, for example, drugs are transported from intestinal epithelial cells into the portal vein and from there to the liver, then through hepatic veins to the heart for general distribution in the bloodstream.

Only a small proportion of a given drug dose will be responsible for producing a specific pharmacological effect—most of the drug will be broken down by metabolizing enzymes (principally within the liver) or excreted unchanged. Although some drug transport is passive, many of the different drug-handling events require dedicated proteins whose function is to transport specific drugs into or out of cells. Within the target tissue are receptor proteins and other proteins in a biological pathway that the drug interacts with to produce the pharmacological effect.

Genetic factors are implicated in the variation between individuals in drug metabolism at each of the different levels: drug absorption (differences in the ability to transport a drug into the bloodstream); drug activation (differences in the ability of liver enzymes to convert a prodrug into the active drug); target response (differences in how the targeted process or pathway responds to a given local concentration of the drug); and catabolism and excretion (differences in the rate at which the drug is catabolized and disposed of).

Phase I and phase II reactions in drug metabolism

Drug metabolism is a defense mechanism: it facilitates excretion of the parent drug and its metabolites, and so limits their ability to accumulate within the body and cause dose-dependent toxicity. It is mostly carried out in the liver (which has multiple enzymes that are responsible for detoxifying drugs and assisting in their excretion), but significant drug metabolism also occurs in some other sites, such as the intestines and kidneys. Small molecule drugs are based on hydrocarbon backbones and so are lipophilic, but drug metabolism allows them to be converted into hydrophilic forms that are easier to excrete from the body.

Phase I reactions are usually carried out by monooxygenases; these work by adding an oxygen atom from molecular oxygen to produce a more polar substance (**Figure 9.7**). Often a hydroxyl group is introduced, or a bulky alkyl group bound to a nitrogen, sulfur, or oxygen atom is replaced by a hydrogen atom. The drug derivative is typically left with a more reactive group, a molecular “handle” that makes it easier for a secondary reaction to be carried out (see below). Sometimes, a phase I reaction also results in *drug activation,* converting an inactive form of a drug, a **prodrug**, into an active drug. For example, the painkiller codeine (methyl-morphine) is a prodrug that is converted by a phase I enzyme in the liver into an active form, morphine.

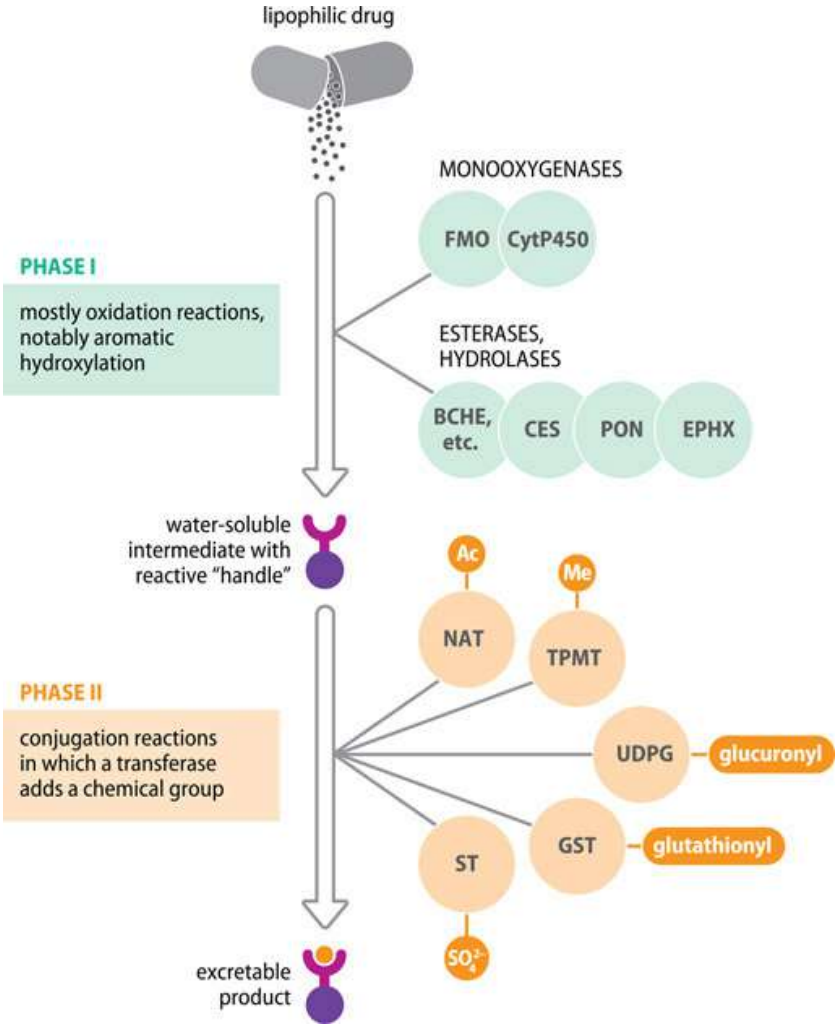


Figure 9.7 Two major stages in drug metabolism. Phase I drug reactions typically result in a more polar drug derivative with a reactive group, a molecular "handle" that makes it easier to accept a chemical group donated by a phase II enzyme. Phase I enzymes are often monooxygenases, notably cytochrome P450 enzymes (CytP450), but also include various other enzymes, notably esterases and other hydrolases. In phase II drug metabolism one of a variety of different transferase enzymes adds a chemical group that facilitates excretion. Note that the sequence shown here occurs commonly, but phase II reactions can sometimes occur without a previous phase I reaction. BCHE, butyrylcholinesterases; CES, carboxyesterases; EPHX, epoxide hydrolases; FMO, flavin-containing monooxygenases; GST, glutathione *S*-transferase; NAT, *N*-acetyltransferase; PON, paraoxonases; TPMT, thiopurine methyltransferase; UDPG, UDP glucuronosyltransferases; ST, sulfotransferases.

Phase II reactions are conjugation reactions, catalyzed by transferases that add one of a variety of chemical groups. As illustrated in Figure 9.7, phase II reactions commonly occur after phase I reactions have introduced a molecular handle for attaching the secondary chemical group. A hydroxyl group attached during phase I, for example, provides a

convenient site for an acetyl group or a sugar (glucuronyl) group to be attached by a phase II enzyme, detoxifying the drug and assisting in its excretion.

Phenotype differences arising from genetic variation in drug metabolism

All drugs work optimally within a certain **therapeutic window**, a range of concentrations within which the therapeutic benefit is optimal without posing any great risks to health. If the concentration is below this range, the therapeutic benefit might be insufficient (drug underdose); if above this range, there is an increasing risk of toxicity (drug overdose).

When drugs are detoxified by metabolism, the concentration of active drug falls; repeated drug doses are required to maintain drug concentrations within the safe therapeutic range. The speed at which a drug is metabolized has consequences for both *drug efficacy*—the degree to which the drug gives therapeutic benefit—and safety. Individuals who eliminate or inactivate drugs comparatively slowly ("slow metabolizers") will have a longer or stronger response to a given concentration of the drug than fast metabolizers will. They can be at risk of a drug overdose if given the usual dose (**Figure 9.8**). They may also be more at risk of adverse reactions if breakdown products from the drug are toxic. Ultrafast metabolizers might gain little therapeutic benefit from a drug (see Figure 9.8).

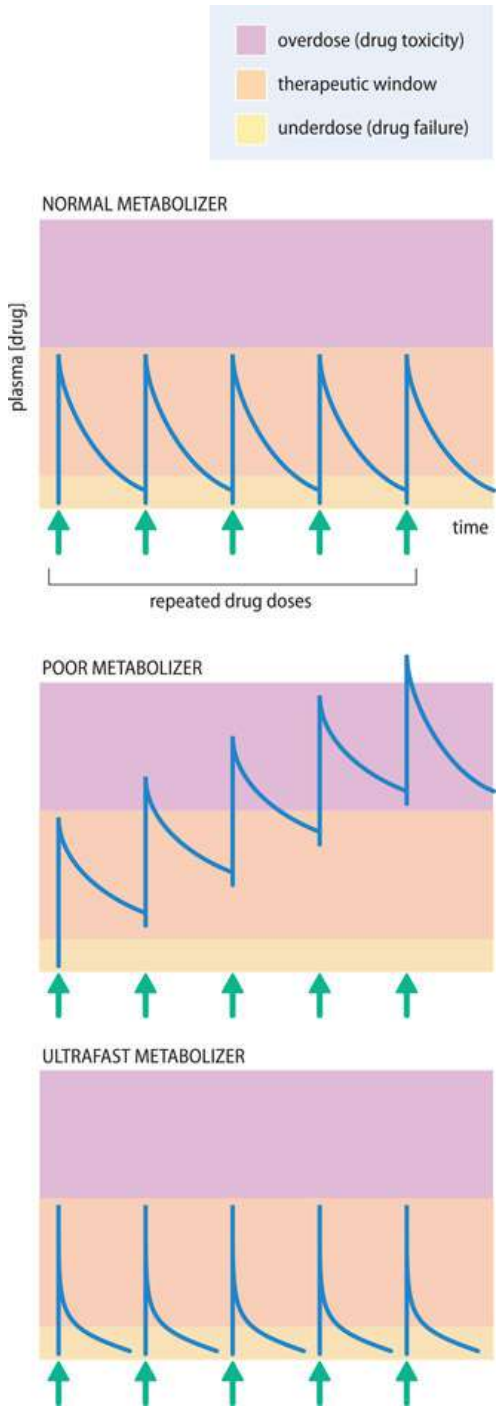


Figure 9.8 The effect of different drug metabolizing rates on plasma drug concentration. The *therapeutic window* is the range of plasma drug concentrations that are of therapeutic benefit without causing extra safety risks due to drug toxicity. Normal metabolizers are expected to benefit from having drug concentrations in the therapeutic window for long periods. If given the normal drug concentration, poor metabolizers can suffer from an overdose: the failure to metabolize the drug quickly means that the drug concentration progressively increases to very high, unsafe levels after repeated doses. Ultrafast metabolizers might gain little therapeutic

benefit, because the drug is rapidly cleared from the plasma after each drug dose. See Figure 9.9 for specific examples of different classes of drug metabolism.

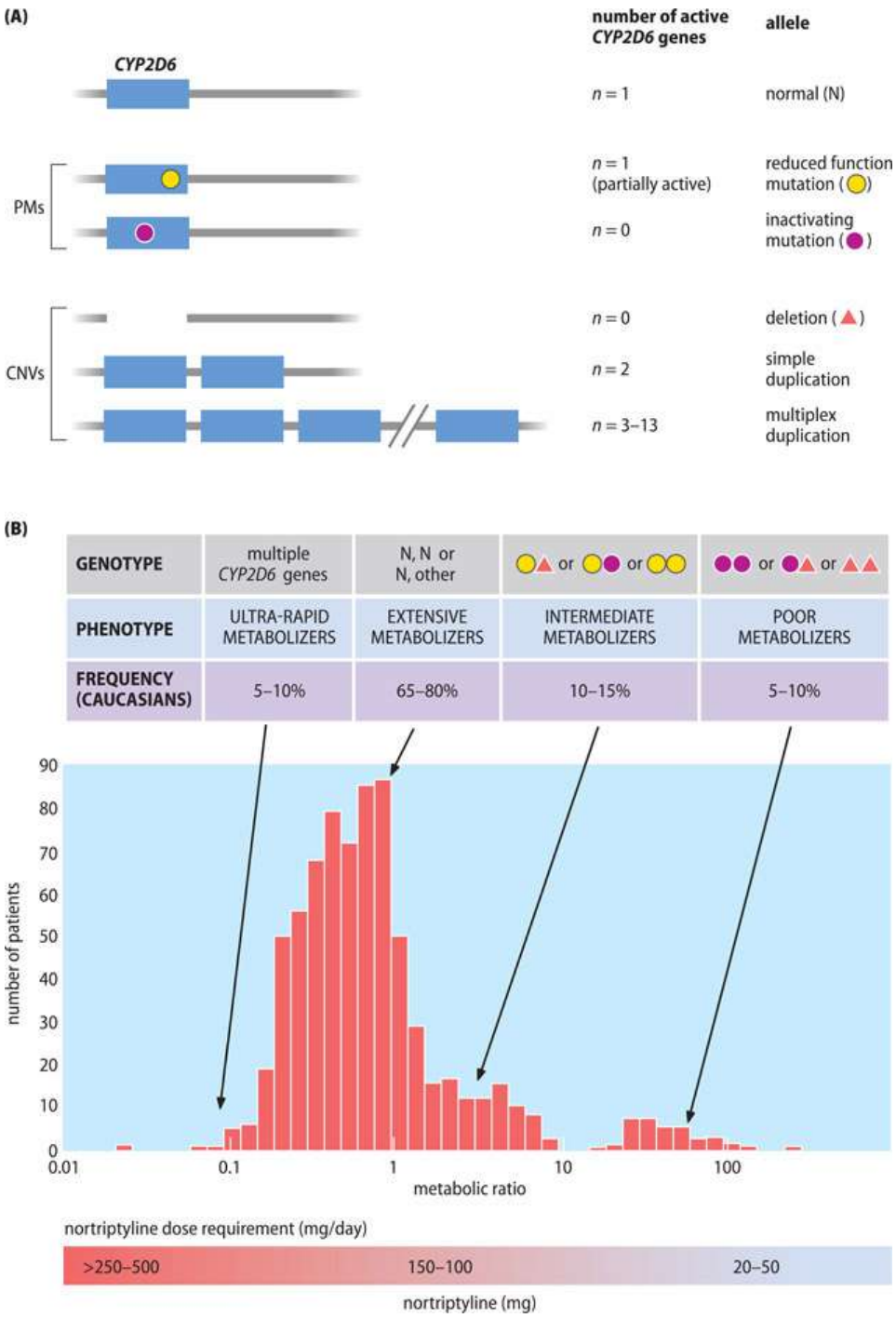


Figure 9.9 CYP2D6 allele classes and correlation between genotypes and drug-metabolizing abilities. (A) *CYP2D6* allele classes. Variation includes both common point mutations (PMs) and copy number variants (CNVs). Multiple *CYP2D6* genes are quite common in some populations, most probably as a result of natural selection (much like the gene amplification giving rise to multiple α-amylase genes in some populations, as

described in Figure 4.8). (B) *CYP2D6* genotypes and drug-metabolizing ability. The histogram in the middle panel shows a range of CYP2D6's drug-metabolizing ability (metabolic ratio) in a study group. To assay CYP2D6 activity, a urinary metabolic ratio is used: after a standard drug dose, the urinary concentration of the substrate drug (debrisoquine in this case) is measured and divided by that of its metabolic product. High ratios show poor conversion as a result of low enzyme activity. The upper panel shows the four classes of metabolizer, their frequencies (in Caucasian populations) and how phenotype relates to genotype. Low metabolizers are at risk of drug overdose and should be given lower drug doses. As an example, the lower panel shows recommended doses of the CYP2D6-metabolized antidepressant nortriptyline, arranged on a sliding scale according to the metabolic ratio shown above. (Adapted from Meyer UA [2004] *Nature Rev Genet* 5:669–676; PMID 15372089. With permission from Macmillan Publishers Ltd.)

Individual drugs can be metabolized by multiple different enzymes encoded by different genes; as well as genetic variation in different genes, environmental factors contribute significantly to how a drug is metabolized. As a result, drug response phenotypes are often multifactorial. Sometimes, however, a specific drug might be metabolized by just one enzyme; genetic variation in the gene making that enzyme can have a predominant contribution to the phenotype—we provide examples in the next sections.

Genetic variation in cytochrome P450 enzymes in phase I drug metabolism

The great majority of phase I drug reactions are carried out by monooxgenases, notably heme-containing enzymes in the cytochrome P450 superfamily (they have in common a spectral absorption peak at 450 nm). Cytochrome P450 enzymes catalyze different types of reaction. In addition to specific drugs, their substrates can be endogenous chemicals, notably certain steroids, and xenobiotics in our food and in the environment.

We have >110 different cytochrome P450 genes, classified into 18 families and multiple subfamilies. Thus, for example, the *CYP2C19* gene gets its name from cytochrome P450 family 2, subfamily C, polypeptide 19. Six cytochrome P450 enzymes catalyze 90 % of the phase I reactions on commonly used drugs. CYP3A4 is involved in metabolizing ~40 % of all drugs; CYP2D6 is another prolific drug-handler. Individual drugs are also often substrates for more than one P450 enzyme. The antidepressant amitriptyline, for example, can be metabolized by each of CYP1A2, CYP2C19, and CYP2D6.

Specific cytochrome P450 enzymes can be induced or inhibited by certain drugs. That can result in unexpected interactions between these drugs and those that are substrates of the enzyme. Comprehensive drug interaction lists are maintained in the Drug Interactions Flockhart Table (available at https://drug-interactions.medicine.iu.edu/MainTable.aspx).

The wide and overlapping specificities of cytochrome P450 enzymes mean that it can often be difficult to correlate how a person metabolizes a specific drug with the activity of any one P450 enzyme. Nevertheless, some drugs are metabolized by just one P450 enzyme, and DNA variation in just a single gene can result in wide differences between people in how they

metabolize the drug. Often the variation is due to simple mutations that change key amino acids or that inactivate gene expression, but occasionally excess gene activity occurs when there are multiple copies of the same cytochrome P450 gene.

Genetic variation in CYP2D6 and its consequences

The *CYP2D6* gene has more than 100 genetic variants, with a continuum of enzyme activity; it is a prime example of how different types of genetic variation in a single enzyme can have a marked effect on the metabolism of certain drugs. As a result of severe inactivating mutations or deletions in both *CYP2D6* alleles, rare individuals have a very low activity of the enzyme. When treated with certain drugs normally metabolized by this enzyme alone, they fail to metabolize and excrete the drug (with high plasma levels for the drugs and low levels of expected catabolic products in urine samples).

Depending on their CYP2D6 activity, people vary in their ability to metabolize drugs for which this enzyme has the predominant role). Four classes of metabolizers are recognized: poor metabolizers (who lack normal alleles, and are comparatively frequent in Caucasian populations); intermediate metabolizers; extensive metabolizers (with one or two active *CYP2D6* alleles); and ultrafast metabolizers (having multiple *CYP2D6* genes as a result of gene amplification) (**Figure 9.9**).

People with very low CYP2D6 activity can show unusually marked sensitivity to certain drugs; they are also at risk of drug overdose if prescribed with normal doses of certain beta-blockers and tricyclic antidepressants. Because CYP2D6 is also the enzyme that converts codeine to morphine, people with very low CYP2D6 activity also get minimal painkilling benefit from codeine. Ultrafast metabolizers may also get little benefit from drugs principally metabolized by CYP2D6 (because they metabolize and detoxify the drugs so quickly).

Genetic variation in other cytochrome P450 enzymes

CYP3A4 activity in the liver shows extensive variability between individuals, but unlike for CYP2D6 there are very few coding sequence variants. And unlike CYP2D6, CYP3A4 is inducible; regulatory mutations are thought to be signifi-cant contributors to the variability.

CYP3A4 is highly related to CYP3A5 (and to CYP3A7, which is normally expressed only at fetal stages); the drug-metabolizing activities of CYP3A4 and CYP3A5 strongly overlap, complicating matters. However, CYP3A5 is less biologically active than CYP3A4: in Caucasian populations *CYP3A5* null alleles predominate and only 10 % of the alleles make an active enzyme.

CYP2C9 deficiency is important in metabolizing the anticoagulant warfarin (as described below), and also results in an exaggerated response to tolbutamide, a hypoglycemic agent used in treating type 2 diabetes. CYP2C19 deficiency shows marked differences between

different ethnic groups, as revealed by the frequency of poor metabolizers (Table 9.2); poor metabolizers require a lower dose of certain drugs such as clopidogrel, an antiplatelet agent that is used to inhibit blood clotting in some diseases such as coronary artery disease.

TABLE 9.2 SIGNIFICANT POPULATION DIFFERENCES IN THE FREQUENCY OF POOR METABOLIZERS OF CYP2D6 AND CYP2C19

Population origin	Frequency (%) of CYP2D6 poor metabolizers	Frequency (%) of CYP2C19 poor metabolizers
Amerindian	0	2.0
Caucasian	7.2	2.9
East and South East Asian*	0.5	15.7
Middle East and North African	1.5	2.0
Polynesian	1.0	13.6
Indian subcontinent (Sri Lankan)	0	17.6
Sub-Saharan African	3.4	4.0

* Excluding the Indian subcontinent. (Data from Burroughs VJ, Maxey RW & Levy RA [2002] *J Natl Med Assoc* 94(10 Suppl): 1–26; PMID 12401060.)

Genetic variation in enzymes that work in phase II drug metabolism

Aromatic *N*-acetylation is a frequent type of phase II drug metabolism. Two types of *N*-acetyltransferase, NAT1 and NAT2, deal with different sets of drugs. Whereas NAT1 is comparatively invariant, NAT2 is polymorphic in a wide range of human populations, with rapid acetylators (who eliminate drugs rapidly) and slow acetylators (who have low NAT2 levels).

Variation in acetylating abilities of NAT2 can be clearly seen as a bimodal distribution using the drug isoniazid, which is used to treat tuberculosis (Figure 9.10). The proportion of slow acetylators is highly variable between different ethnic groups but can be very high in some populations, notably in Caucasian populations (Table 9.3).

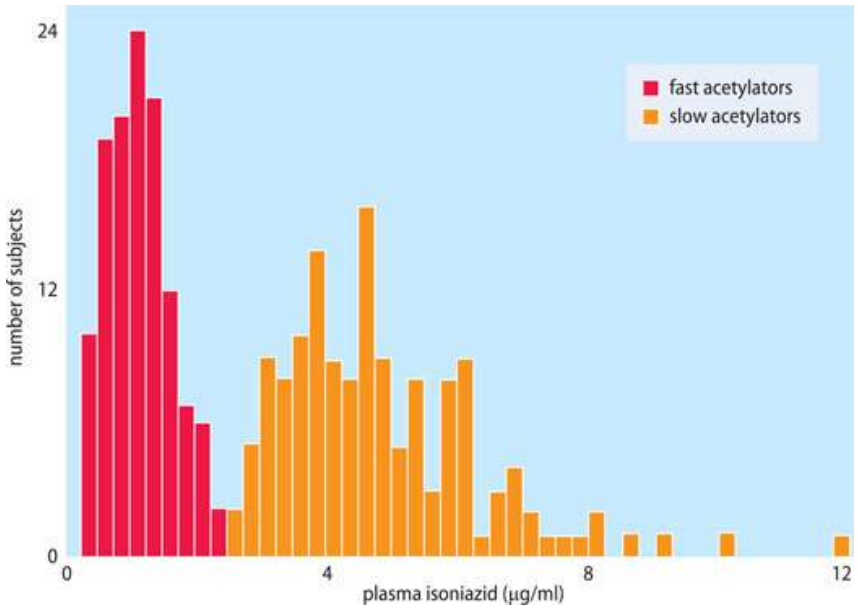


Figure 9.10 A bimodal distribution of plasma levels of isoniazid as a result of genetic polymorphism in the *NAT2* (*N*-acetyltransferase 2) gene. Plasma concentrations were measured in 267 normal subjects 6 hours after an oral dose of isoniazid. Fast acetylators removed the drug rapidly. The number of slow acetylators (presumed to be homozygotes or compound heterozygotes for severe inactivating mutations) was almost the same as the number of fast acetylators. This suggests that only about 30 % of *NAT2* alleles in the study group were producing active enzyme. (Adapted from Price Evans DA, Manley KA & McKusick VA [1960] *Br Med J* ii:485–491. With permission from BMJ Publishing Group Ltd.)

TABLE 9.3 SIGNIFICANT POPULATION DIFFERENCES IN THE FREQUENCY OF NAT2 SLOW ACETYLATORS

Population origin	Frequency (%) of NAT2 slow acetylators
Caucasian	58
Chinese	22
Eskimo	6
Japanese	10
Sub-Saharan African	51

Data from Wood AJ & Zhou HH [1991] *Clin Pharmacokinet* 20:350–373; PMID 1879095.)

The slow acetylators take longer to eliminate drugs (and other xenobiotics) and so often show enhanced sensitivity to drugs metabolized by NAT2; they also appear to be more susceptible to certain cancers, notably bladder cancer. Nevertheless, natural selection appears to have driven an increase in frequency of the slow-acetylator phenotype in some populations. A possible explanation is that certain chemicals in well-cooked meat are converted by NAT2 into carcinogens; individuals with slow-acetylator phenotypes would be comparatively protected in populations that have had a long tradition of eating well-cooked meat.

Variation in other phase II enzymes is also significant. Polymorphism in thiopurine methyltransferase is a particular clinical concern when using certain immunosuppressant drugs such as 6-mercaptopurine, which is commonly used to treat childhood leukemia. (About 1 in 300 children do not express thiopurine methyltransferase; for them 6-mercaptopurine is toxic.) The glutathione *S*-transferase (GST) superfamily includes some enzymes such as GSTM1 and GSTT1 that are encoded by genes susceptible to gene deletion via unequal crossover. As a result, inactive alleles are very common (about 50 % of the *GSTM1* alleles in people of northern European ancestry are gene deletions, for example). People with consequently low levels of these enzymes find it difficult to cope with high doses of drugs that are processed by them.

The UDP glucuronosyltransferase superfamily includes polymorphic enzymes involved in metabolizing different substrates, including drugs used in cancer chemotherapy. For example, the prodrug irinotecan is converted into an active anti-tumor form in the liver (it inhibits DNA topoisomerase, an enzyme needed for DNA replication) and is normally processed by the enzyme UGT1A1. The common *UGT1A1*28* promoter polymorphism results in reduced production of this enzyme, and is frequent in many populations (1 in 3 people are homozygotes in sub-Saharan Africa, and 1 in 5 in the Indian subcontinent). The **28/*28* homozygotes have a much higher risk of serious bone marrow and gastrointestinal toxicity.

Altered drug responses resulting from genetic variation in drug targets

In the sections above we have considered genetic variation in enzymes performing phase I and phase II drug metabolism, and how that variation affects drug pharmacokinetics (including how fast a drug is metabolized and excreted). In this section we consider how genetic variation affects the pharmacodynamics of drugs.

The efficacy of a drug is partly determined by genetic variation in *drug targets,* the molecules that the drug must interact with in the cells of the target tissue. Drug targets will typically include receptors, signaling molecules, and other molecular components of biological pathways that the drug interacts with to have its pharmacogenetic effect.

Genes encoding drug receptors quite often show polymorphisms or variants that lead to clinically significant altered responses to drugs. Examples of clinically significant genetic variation in drug receptors include variants of beta-adrenergic receptors, cell surface receptors that have central roles in the sympathetic nervous system. Two of these receptors, ADRB1 and ADRB2, are widely used as drug targets in therapeutic approaches for various common and important diseases including asthma, hypertension, and heart failure. Genetic variation in both ADRB1 and ADRB2 has been linked to altered responses to drugs. Other examples include variation in the H2RA serotonin receptor and the RYR1 ryanodine receptor (**Table 9.4**).

TABLE 9.4 EXAMPLES OF HOW GENETIC VARIATION IN DRUG TARGETS CAUSE ALTERED RESPONSES TO THERAPEUTIC DRUGS

Drug target	Function	Polymorphism or variant	Example of drug treatment	Effect of polymorphism or variant
ACE	angiotensin-converting enzyme	Alu repeat indel polymorphism in intron of *ACE* gene	use of ACE inhibitors—captopril and enalapril—to treat heart failure	drugs are more effective in Alu^-/Alu^- homozygotes
ADRB1	β_1 adrenergic receptor	common R389G polymorphism	beta-blockers, such as bucindol, for reducing heart disease risk	reduced cardiovascular response to drugs
ADRB2	β_2 adrenergic receptor	common R16G polymorphism	albuterol for treating asthma	homozygotes are much less likely to respond to treatment*
RYR1	ryanodine receptor	different mutations	inhalation anesthetics	potentially fatal (see Clinical Box 11)

* Note: the R16G polymorphism has significant effects on short-acting agonists but little effect on long-acting agonists, which now constitute the more widely used treatment.

Some therapeutic drugs are designed to specifically inhibit key enzymes that have pivotal roles in biological pathways underlying common diseases. Examples include statins, which were designed to inhibit HMG CoA reductase (lowering cholesterol levels, and so reducing blood pressure and the risk of cardiovascular disease); warfarin, an inhibitor of a key enzyme in the maturation of several blood-clotting factors (see the next section); and inhibitors of angiotensin-converting enzyme (ACE). The last enzyme is also an important regulator of blood pressure (and several other functions), and an insertion/deletion polymorphism due to the variable presence of an Alu repeat in intron 15 of the *ACE* gene is associated with variation in ACE activity. People who are homozygous for the Alu deletion allele have about twice the enzyme activity of those who have two Alu insertion alleles; this difference is thought to be an important contributor to variable responses to ACE inhibitors (see Table 9.4).

We give a summary of serious adverse drug reactions in **Clinical Box 11**. In most cases, the genetic variation associated with the effect is primarily confined to a single locus. However, for many drugs genetic variation at multiple loci might be important. One common example where we know some of the details concerns treatment with the anticoagulant warfarin.

CLINICAL BOX 11 WHEN PRESCRIBED DRUGS CAN BE DANGEROUS AND SOMETIMES DEADLY, DEPENDING ON A PATIENT'S GENOTYPE

Therapeutic drugs and other drugs administered in medical procedures (such as anesthetics) can produce extreme responses in some people. Adverse drug reactions are very common, being responsible for a signifi-cant proportion of all hospital admissions (nearly 7 % in a UK

study) and can result in disability or permanent damage, birth defects, and an extraordinary number of fatalities (about 100 000 deaths each year in the US alone).

Adverse drug reactions have various causes. Type A reactions are relatively common and dose-dependent; they are predictable from the drug pharmacology and are usually mild. Type B reactions are idiosyncratic reactions that are not related simply to drug dose; they are rare but can often be severe (Table 1). Genetic variants are important in both types of reaction.

TABLE 1 SOME CLASSES OF SEVERE TYPE B ADVERSE DRUG REACTIONS

Drug-induced injury or toxicity	Associated drugs (examples)	Comments and examples
Apnea (respiratory paralysis)	suxamethonium (succinyl choline)	this drug works as a fast-acting muscle relaxant and is used before surgery. Normally the effects of the drug wear off quite quickly when the drug is metabolized by the enzyme butylcholinesterase. Low metabolizers are at risk of apnea—they remain paralyzed and unable to breathe after surgery because they cannot regain their muscle function quickly enough and may require extended ventilation
Prolongation of cardiac QT interval	thioridazine, clarithromycin, terfenadine	induced by many different drugs. Associated with polymorphic ventricular tachycardia, or torsades de pointes (see Figure 1), which can be fatal
Hematologic toxicity	6-mercaptopurine, azathioprine	thiopurine S-methyltransferase (TPMT) inactivates these immunosuppressant drugs by adding a methyl group. In people with two low-activity TPMT alleles, the drugs are metabolized slowly; if normal doses are given, the drugs accumulate and can result in life-threatening bone marrow toxicity
Hemorrhage	warfarin	see text
Hypersensitivity reactions	abacavir, carbamazepine, allopurinol	inappropriate immune reactions to otherwise nontoxic drugs can have broad manifestations. When treated with the anti-HIV drug abacavir, about 5 % of patients demonstrate skin, gastrointestinal, and respiratory hypersensitivity reactions that can sometimes be fatal. Treatment with the anticonvulsant carbamazepine or allopurinol (used in treating gout) can induce cutaneous adverse drug reactions, including toxic epidermal necrolysis

Drug-induced injury or toxicity	Associated drugs (examples)	Comments and examples
Liver injury	flucloxacillin, isoniazid, allopurinol	individuals with certain HLA antigens are at increased risk of induced liver disease for some drugs, such as *HLA-B*5701* in the case of flucloxacillin
Muscle toxicity	halothane, isoflurane, various statins	statins and several other drugs are associated with usually mild myopathies. Sometimes, however, more severe cases show rhabdomyolysis (breakdown of muscle tissue) that can result in death. In response to inhalation anesthetics (halothane, isoflurane), individuals who have inactivating mutations in the ryanodine receptor gene develop life-threatening rhabdomyolysis and an extreme rise in temperature, a form of malignant hyperthermia (OMIM 145600)

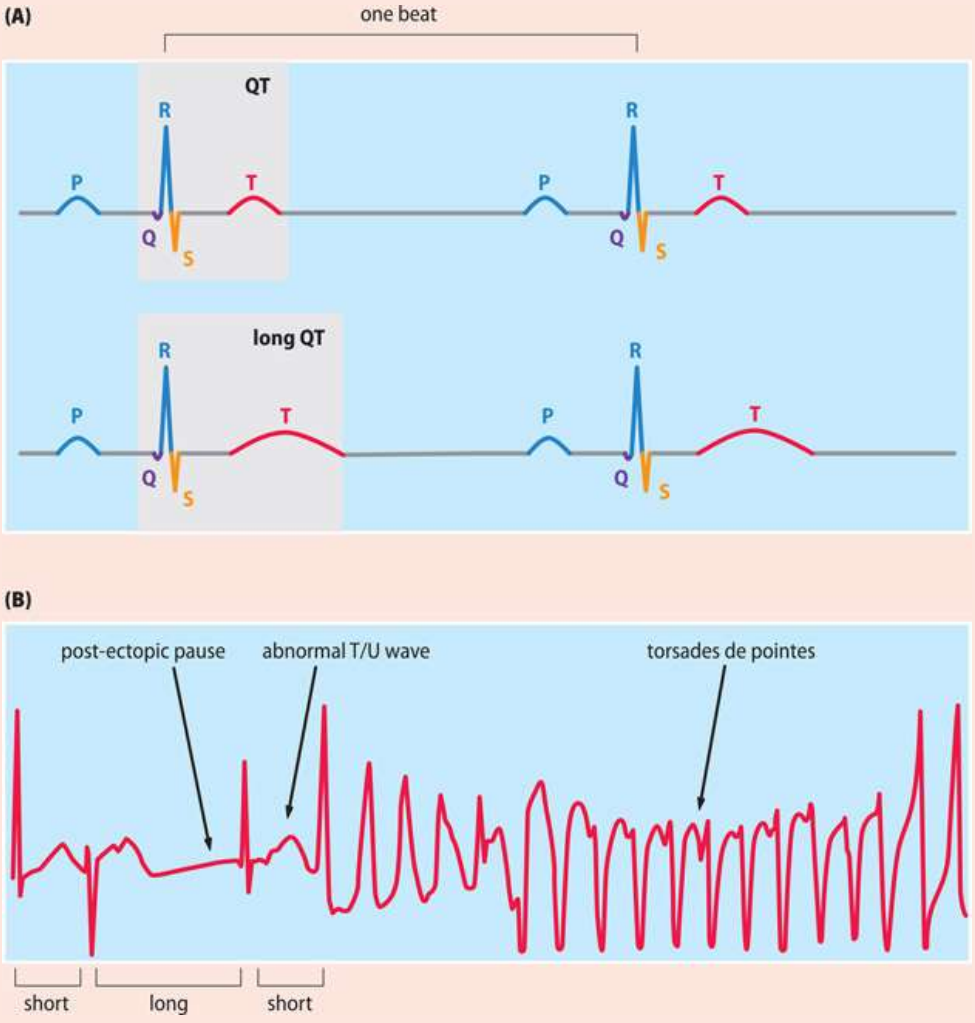


Figure 1 Drug-induced prolongation of the cardiac QT interval and torsades de pointes. (A) Cardiac depolarization–repolarization cycle. Specific repeated features are labeled from P to T. The *QT interval,* the shaded interval that spans the onset of the QRS complex until the end of the T wave, represents the time taken for one complete cycle of ventricular depolarization and repolarization. Certain drugs can prolong the QT interval *(long QT*), and this can sometimes induce a rapid beating of the heart, which often manifests itself as torsades de pointes (TdP). (B) Cardiac rhythm profile in a patient with drug-induced TdP. Notice the short–long–short initiating ventricular cycle, pause-dependent long QT interval, and abnormal TU wave preceding the development of TdP. This type of ventricular arrhythmia can self-terminate, but it can also degenerate into potentially fatal arrythmias such as ventricular fibrillation. (B, Adapted from Yap YG & Camm AJ [2003] *Heart* 89:1363–1372; PMID 14594906. With permission from BMJ Publishing Group Ltd.)

When genotypes at multiple loci in patients are important in drug treatment: the example of warfarin

Warfarin is prescribed for patients at risk of developing clots within blood vessels (thrombosis), including clotting that can block arteries (embolism). Delivering the optimal warfarin dosage is clinically very important because there is a narrow therapeutic window: if the administered warfarin level is too low, the patient remains at risk of thrombosis and embolism; if it is too high, there is a risk of life-threatening hemorrhage. The final warfarin dose is critical, but because of genetic variation the optimal dose varies enormously between individuals.

Chemically, warfarin is a mixture of two isomers. The (*S*)-warfarin isomer is three to five times as potent as the (*R*)-isomer, has a shorter half-life, and is metabolized predominantly by CYP2C9. Two common polymorphisms in CYP2C9 result in a large decrease in enzyme activity, down to 12 % of wild-type activity for CYP2C9*1, and to just 5 % of normal enzyme activity for CYP2C9*3. Patients with one or two of these common alleles are at increased risk of hemorrhage (presumably because drug metabolism takes longer). It soon became clear, however, that these polymorphisms could explain only a part of the genetic variation in the final warfarin dose.

In 2004 the drug target of warfarin was found to be vitamin K epoxide reductase complex subunit 1 (VKORC1), an enzyme that converts oxidized vitamin K to its reduced form; afterward, association studies showed that genetic variation in VKORC1 was associated with variation in the final warfarin dose. Vitamin K is an indispensable cofactor for the enzyme that converts inactive clotting proteins to give four of our blood clotting factors (factors II, VII, IX, and X). By inhibiting vitamin K epoxide reductase, warfarin inhibits the recycling of vitamin K; the consequent decreased supply of vitamin K inhibits the formation of these four clotting factors.

Subsequently a genomewide association study also implicated a common V433M polymorphism in CYP4F2, a cytochrome P450 enzyme that works as a vitamin K oxidase

(Figure 9.11). Variations in VKORC1 and CYP2C9 remain the largest genetic determinants, accounting for about 40 % of the variation in the final warfarin dose. However, other factors, such as aging and the simultaneous administration of other medicines, also have a very significant influence on the dose of warfarin that is prescribed.

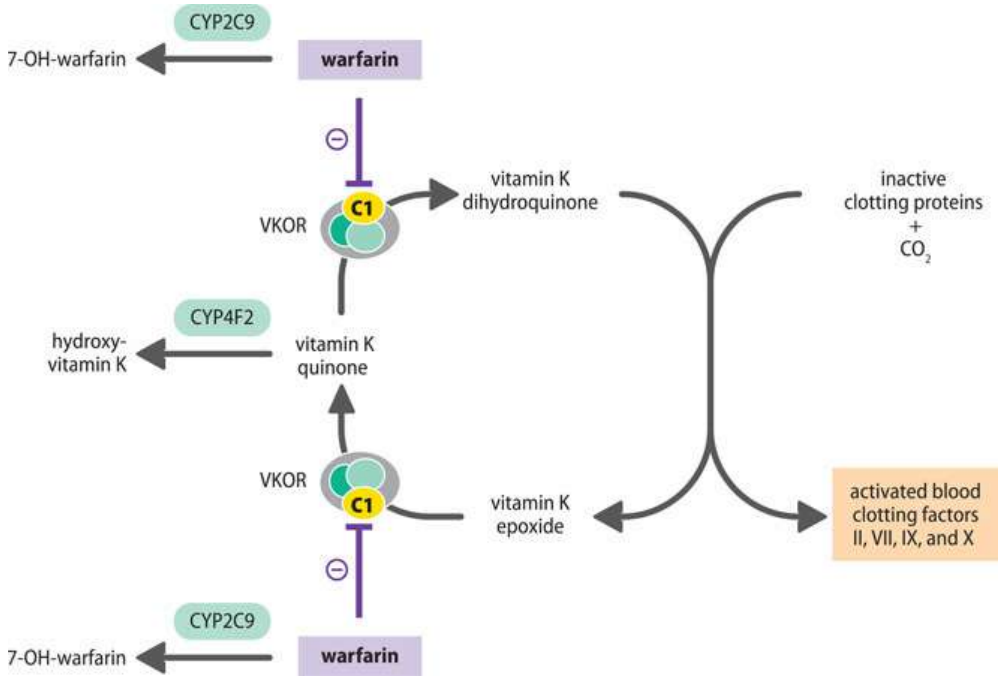


Figure 9.11 Roles of warfarin in anticoagulation and genetic variants affecting final warfarin drug dose. Warfarin is a therapeutic anticoagulant prescribed for people at risk of thrombosis and embolism. It works by inhibiting VKORC1, the C1 subunit of the vitamin K epoxide reductase complex (VKOR). VKOR is a precursor of vitamin K dihydroquinone, which activates four blood clotting factors: factors II, VII, IX, and X. When VKORC1 is inhibited, the supply of vitamin K is decreased, resulting in a reduced supply of activated clotting factors II, VII, IX, and X. In addition to VKORC1, genetic variation in at least two cytochrome P450 enzymes is known to be associated with variation in the final warfarin dose needed: CYP2C9 converts warfarin to an inactive form, 7-hydroxywarfarin, and CYP4F2 metabolizes vitamin K quinone.

Translating genetic advances: from identifying novel disease genes to therapeutic small molecule drugs

The quest to define the molecular basis of genetic disorders has identified many previously unknown genes, often after labor-intensive studies. Time-consuming studies are then needed to work out how the genes work normally, and in disease states (as mimicked using cultured cells or animal models). Thereafter, therapies can be developed to tackle the *cause* of the disease rather than just the symptoms.

As we will see later in this chapter, gene therapy is becoming an attractive option for treating the cause of several genetic diseases (notably recessive monogenic disorders, as described later in this chapter). An alternative possibility, however, is to develop therapeutic

small molecule drugs that possess aromatic hydrocarbon backbones (just like the familiar aspirin, codeine, and so on that we have long been used to).

The pharmaceutical industry is the environment for developing therapeutic small molecule drugs. Assays are devised in which the fault caused by the mutant gene is replicated in cultured cells (or sometimes even in very simple model organisms in the case of highly conserved genes), and then high throughput screening is carried out by exposing suitable mutant cells to tens of thousands of different synthetic small molecules in parallel. The aim is to identify individual small molecules that somehow overcome the adverse effects caused by pathogenic mutation; they then become candidate therapeutic drugs for further detailed testing. We provide three examples below to illustrate how identifying novel genes for monogenic disorders has led to developing novel therapeutic small molecule drugs.

The first example is an early success story in which genetic advances led to a promising drug target that was screened to identify a new class of very valuable drugs. The second illustrates treating a common recessive disorder which, however, is not so amenable to gene therapy, and the need for stratifying mutations according to their functional effects in cells. In the third example, working out what the disease normally does and discovering that the defect involved a specific molecular pathway allowed the application of a previously identified drug that works in the same pathway and has largely replaced a surgical treatment method.

Familial hypercholesterolemia: new and valuable drugs

With a frequency of more than 0.2 % in most populations, familial hypercholesterolemia (OMIM 143890), an autosomal dominant disorder, is the most common single-gene disorder. Affected persons have extremely high cholesterol levels, irrespective of diet. Most cases are due to mutations in the low-density lipoprotein receptor gene, *LDLR*. Heterozygotes typically develop coronary artery disease in the fourth or fifth decade; rare homozygotes are much more severely affected, and most suffer a heart attack before the age of 20 years.

LDLR imports cholesterol-containing low-density lipoprotein into liver cells, where it represses cholesterol synthesis as part of a homeostatic mechanism. An *LDLR* loss-of-function mutation results in less LDLR being made, resulting in an increase in endogenous cholesterol synthesis. Hydroxymethylglutaryl (HMG) CoA reductase was known to be the rate-limiting enzyme in the endogenous cholesterol biosynthesis pathway, and so represented a very promising drug target.

Screening for hydrocarbon-based small molecules that inhibit HMG CoA reductase identified statins, a class of drugs effective in lowering cholesterol. Drugs such as these have been hugely important, being widely prescribed to reduce the general risk of heart disease, not just to treat familial hypercholesterolemia.

Cystic fibrosis: not an easy prospect for therapy

In 1989 a previously unstudied gene was found to be the locus for cystic fibrosis. Named the cystic fibrosis transmembrane regulator gene *(CFTR),* it was subsequently shown to function as an anion channel that allows chloride and bicarbonate ions to pass through the plasma membrane of cells. A great deal was discovered about aspects of CFTR biology but even twenty years after gene discovery, slow progress had been made in understanding the pathogenesis, and new therapies were lacking. In the last decade, however, significant progress has been made.

As we explain in Section 9.4, cystic fibrosis is an example of a recessive disorder that is difficult to treat by gene therapy, prompting alternative efforts to develop therapeutic small molecule drugs. Different drugs may be applicable, according to their effect at different functional levels (see **Figure 9.12** for one scheme that classifies *CFTR* mutations into six categories according to how they affect normal gene function).

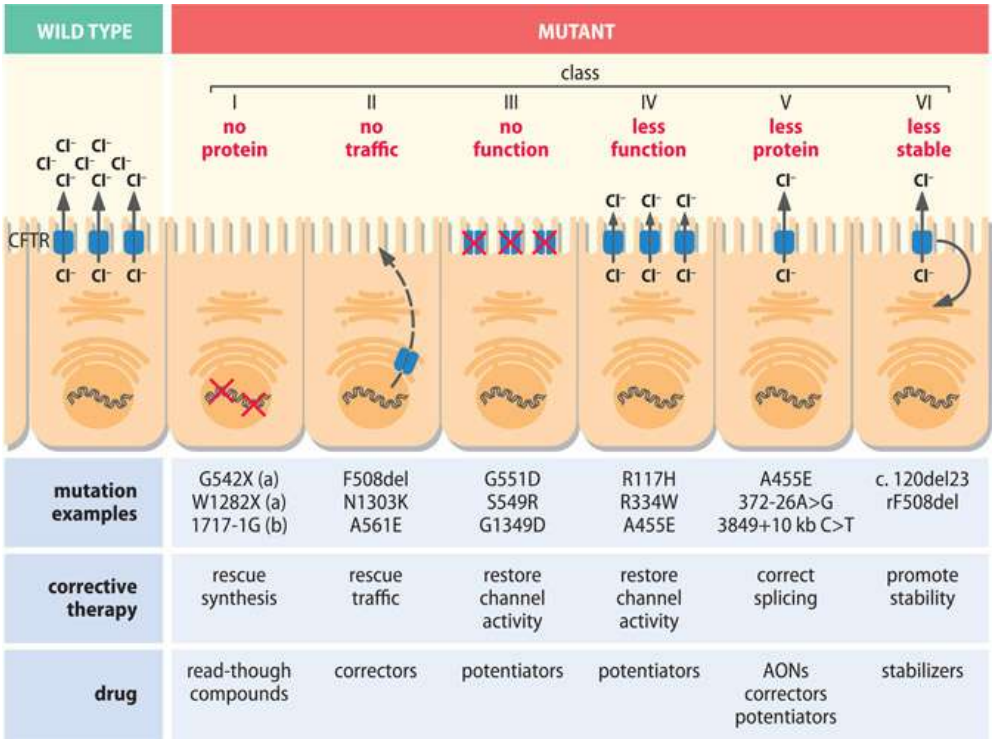


Figure 9.12 Functional classification of *CFTR* mutations causing cystic fibrosis and major targeted CFTR modulator therapies and drug type nomenclature. For mutation classes 1 and 2, the CFTR protein isn't produced or doesn't get to the apical membrane because of problems with protein folding or trafficking within the cell. For mutation classes III-VI, the CFTR protein does get incorporated into the apical membrane, but the ion channel doesn't open (III), doesn't conduct ions properly (IV), is present in low amounts (V), or shows instability (VI). See text for drug classes and progress using them. AONs, antisense oligonucleotides; WT, wild type. Adapted with permission from Amaral MD (2015) *J Intl Med* 277:155–166; PMID 25266997.

The bulk of mutations leading to failure to make a CFTR protein result from premature termination codons (nonsense mutations and mutations causing translational frameshifting).

For patients with these mutations a generic small molecule drug may be considered that can suppress the effect of premature termination codons irrespective of the type of disease. We consider such “read-through compounds” later in this section.

To obtain small molecule drugs that specifically target mutant CFTR protein, high throughput screening is required. In principle, the small molecule drugs can work by binding to and modulating the effect of mutant CFTR proteins in different ways as listed below.

- *Correctors* assist mutant CFTR to adopt a suitable conformation, enabling them to move from the endoplasmic reticulum to the plasma membrane.
- *Potentiators* facilitate ion channel gating (opening the channel) and conduction (increasing the flow of ions through an open channel).
- *Stabilizers* help keep the CFTR protein anchored to the plasma membrane.
- *Amplifiers* increase the amount of CFTR protein made.

The first breakthrough came when ivacaftor, a potentiator drug marketed by Vertex Pharmaceuticals, was shown to be effective in treating cystic fibrosis patients with the Gly551Asp (G551D) mutation. In the cells of these patients the mutant CFTR protein causes the ion channel to fail to open, but ivacaftor helps to reopen it. After promising clinical results, ivacaftor received regulatory approval in the US in 2012. Although it was subsequently found to be useful for treating patients with some other minor *CFTR* mutations, however, ivacaftor treatment is appropriate for just 5 % of cystic fibrosis patients.

The predominant cystic fibrosis mutation, accounting for around 70 % of pathogenic *CFTR* mutations in many Caucasian populations, produces the Phe508del (F508del) CFTR mutant protein. The problem here is that the mutant CFTR not only misfolds and gets trapped in the endoplasmic reticulum (where it is then targeted for destruction), but any of the mutant protein that does manage to get to the plasma membrane is functionally inactive. In 2019, however, a couple of phase 3 multicenter clinical trials demonstrated that a combination of two folding correctors (elexacaftor and tezacaftor) and the ivacaftor potentiator were both efficacious and safe in treating cystic fibrosis patients with the Phe508del mutant. The combined drug, called Trikafta, is an effective therapy for patients possessing a *CFTR* allele making the mutant F508 del. variant, representing 90% of cystic fibrosis patients in the USA.

Tuberous sclerosis: from a biological pathway to new treatments

Tuberous sclerosis complex is an autosomal dominant disorder in which benign (noncancerous) tumors develop in many organs and can disrupt how they function (tumors of the central nervous system and kidneys are the leading causes of the morbidity and mortality). Additional abnormalities of cell migration and function in the brain lead to seizures, autism, and learning difficulties. The disorder can be caused by mutations in the *TSC1* or *TSC2* genes that encode components of the TSC1–TSC2 protein complex. Until these novel genes were

identified, nothing was known about the molecular pathogenesis of the disorder, and the only treatment options were the often problematic surgical removal of tumors.

The TSC1–TSC2 complex was then found to be part of the mTORC1 growth signaling pathway. When the TSC1–TSC2 complex is disrupted by mutations in either *TSC1* or *TSC2,* mTORC1 signaling is constitutively active; downstream targets become activated by phosphorylation, driving protein synthesis and cell growth (Figure 9.13). A principal subunit of the mTORC1 complex is the mTOR (mammalian target of rapamycin) protein. Rapamycin, also called sirolimus, is an antifungal antibiotic first isolated from a strain of *Streptomyces* in the 1970s. Both it and a closely related drug, everolimus, are powerful mTOR inhibitors. They had initial applications as immunosuppressants to prevent organ transplant rejection and then as anti-proliferative agents for treating certain cancers.

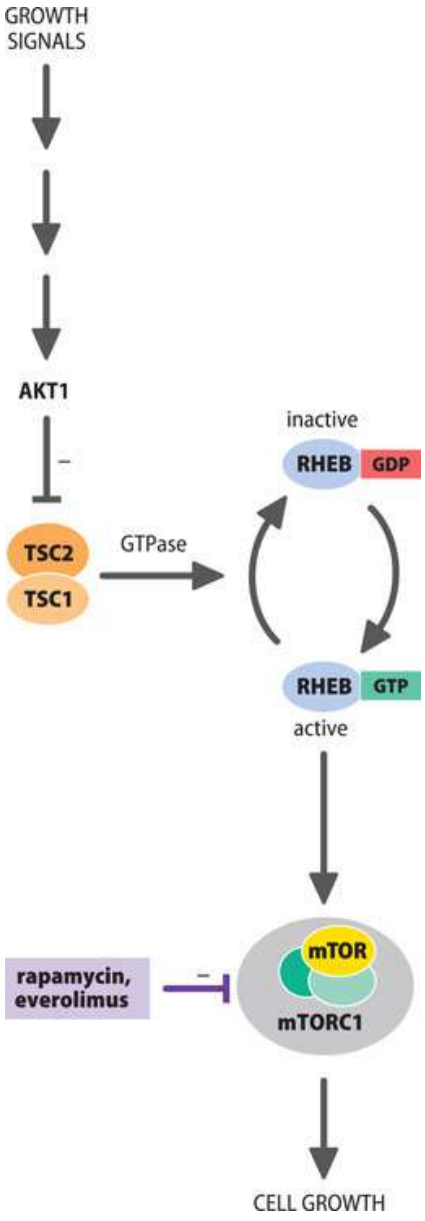


Figure 9.13 Therapeutic targeting of mTORC1 signaling in tuberous sclerosis complex. In tuberous sclerosis complex, the problem is that mutations in either *TSC1* or *TSC2* disrupt the TSC1–TSC2 protein complex. Normally, TSC2 acts as a GTPase and stimulates the formation of an inactive form of the RHEB regulator of the mTORC1 complex (but can be countermanded when growth signals repress TSC1–TSC2). But in tuberous sclerosis complex, the disruption of the TSC1–TSC2 complex causes the RHEB regulator to be activated so that mTORC1 signaling is constitutively active, and growth is no longer regulated as normal. Rapamycin (also called sirolimus) and everolimus, an *O*-(2-hydroxyethyl) derivative of rapamycin, are effective inhibitors of mTOR, a major subunit of the mTORC1 complex, and work to suppress cell growth. AKT1 is also known as protein kinase B.

The tuberous sclerosis complex was discovered to arise from unregulated growth due to abnormal regulation of the mTORC1 pathway. Thereafter various clinical trials were carried out with mTOR inhibitors. Everolimus and related mTOR inhibitors were found to be effective and safe drugs for treating various aspects of tuberous sclerosis, including angiomyolipomas (which are very common in adult Marfan syndrome patients and can lead to renal failure and a need for dialysis) and also subependymal giant cell astrocytomas (SEGAs). Surgical removal used to be the treatment option for managing SEGAs but the surgery can be particularly difficult, and has now largely been replaced by treatment with mTOR inhibitors.

Translating genomic advances and developing generic drugs as a way of overcoming the problem of too few drug targets

Most small molecule drugs work by binding to specific protein targets, blocking their interactions with other molecules. However, only a rather small percentage of protein targets are susceptible to drugs. In many cases, small molecule drugs cannot block interactions between two types of protein because the interacting surfaces of the proteins are too smooth (small molecule drugs are most effective when they can sneak into clefts and pockets within proteins). Some types of protein are easier targets: more than 50 % of drug targets belong to one of four types of protein (class I G-protein-coupled receptors, nuclear receptors, ligand-gated ion channels, and voltage-gated ion channels); protein kinases are another favorite target.

A survey published in 2006 estimated that there were just 324 protein targets for all approved drugs, but by a decade later that number had more than doubled, reaching a total of 667 human protein targets (PMID 27910877). Recently acquired massive data sets generated through genomics, high-throughput transcriptomics, proteomics, and bioinformatics have helped.

Another potential solution to obstacles in identifying novel drug targets is to develop generic drugs not focused on a specific gene product. Because they might have quite widespread applicability, generic drugs might also have the merit of reducing costs (which can be exceptionally high for “orphan drugs” used to treat rare diseases).

A prominent class of generic drugs is made up of "read-through" compounds, small molecule drugs that can suppress stop codons so that translation continues (translational "readthrough"). The potential applications are huge: nonsense mutations are responsible for causing anywhere from 5–70 % of individual cases for most inherited diseases. A recently identified drug, Ataluren (or PTC124), seemed promising: it appeared to cause readthrough of premature nonsense mutations (notably UGA), without affecting the recognition of normal stop codons, and with little evidence of toxicity. However, there have been concerns that PTC 124 activity might be due to off-target effects (PMID 23824517), and its efficacy can be limited (when treated with Ataluren, patients with cystic fibrosis due to *CFTR* nonsense mutations showed just a modest improvement in lung function).

Developing biological drugs: therapeutic proteins produced by genetic engineering

In recent years a new drug class, biological drugs (often called ***biologics***), has been developed. The great majority are therapeutic recombinant proteins, that is, genetically engineered proteins administered to patients. They include genetically engineered antibodies (the subject of the section following this one) and other genetically engineered proteins (described in this section).

Certain genetic disorders resulting from deficiency of a specific protein hormone or blood protein can be treated by administering an external supply of the missing protein. To ensure greater stability and activity, the proteins are often PEGylated, that is, conjugated with PEG [poly(ethylene glycol)]. The increased size of the protein–PEG complex means reduced renal clearance, so that the protein spends more time in the circulation. Pegylation can also make the protein less immunogenic.

Therapeutic proteins were often previously extracted from animal or human sources, but there have been safety issues. A safer alternative is to use therapeutic "recombinant" proteins made by cloning the desired human gene and expressing it to make protein, usually within mammalian cells, such as human fibroblasts or the Chinese hamster ovary cell line. (Mammalian cells are often needed because many proteins undergo post-translational modifications, such as glycosylation, that show species differences in the pattern of modification.) Recombinant human insulin was first marketed in 1982; **Table 9.5** gives several subsequent examples.

TABLE 9.5 EXAMPLES OF THERAPEUTIC RECOMBINANT PROTEINS

Recombinant protein	For treatment of
Insulin	diabetes
Growth hormone	growth hormone deficiency
Blood clotting factors VIII and IX	hemophilia

For genetically engineered therapeutic antibodies, see Table 9.6.

Recombinant protein	For treatment of
Interferon α	hairy cell leukemia; chronic hepatitis
Interferon β	multiple sclerosis
Interferon γ	infections in patients with chronic granulomatous disease
Tissue plasminogen activator	thrombotic disorders
Leptin	obesity
Erythropoietin	anemia

For genetically engineered therapeutic antibodies, see Table 9.6.

TABLE 9.6 EXAMPLES OF LICENSED THERAPEUTIC MONOCLONAL ANTIBODIES (mAbs) FOR TREATING COMMON GENETIC DISEASE

Disease category	Target	mAb generic name (trade name)	Disease treated
Autoimmune or immunological	IgE	omalizumab (Xolair)	asthma
	Integrin α4	natalizumab (Tysabri)	multiple sclerosis; Crohn's disease
	TNFα	certolizumab pegol (Cimzia)	Crohn's disease; rheumatoid arthritis
		adalimumab (Humira)*	
Cancer	EGFR	panitumumab (Vectibix)*	EGFR-expressing metastatic colorectal cancer
	HER2	trastuzumab (Herceptin)	HER2-positive metastatic breast cancer
	VEGF	bevacizumab (Avastin)	colorectal, breast, renal, and NSCL cancer; age-related macular degeneration
Other diseases	PCSK9	alirocumab (Praluent)	hypercholesterolemia that doesn't respond well to statin treatment; atherosclerosis
		evolocumab (Repatha)*	

CD11a, white blood cell antigen; IL2R, interleukin type 2 receptor; IgE, immunoglobulin E; TNFα, tumor necrosis factor α EGFR, epidermal growth factor receptor; HER2, human epidermal growth factor receptor 2; NSCL cancer, non-small-cell lung cancer; VEGF, vascular endothelial growth factor; RSV, respiratory syncytial virus

* Fully human antibodies; all others are humanized antibodies (see Figure 9.14 for an explanation of different monoclonal antibody classes).

Disease category	Target	mAb generic name (trade name)	Disease treated
	VEGF	ranibizumab (Lucentis)	"wet" age-related macular degeneration

CD11a, white blood cell antigen; IL2R, interleukin type 2 receptor; IgE, immunoglobulin E; TNFα, tumor necrosis factor α EGFR, epidermal growth factor receptor; HER2, human epidermal growth factor receptor 2; NSCL cancer, non-small-cell lung cancer; VEGF, vascular endothelial growth factor; RSV, respiratory syncytial virus

* Fully human antibodies; all others are humanized antibodies (see Figure 9.14 for an explanation of different monoclonal antibody classes).

Some human proteins are required in very high therapeutic doses, beyond the production capabilities of cultured cell lines. Transgenic animals such as transgenic sheep or goats are an alternative source; here, the desired protein is secreted in the animal's milk, aiding purification. In 2009, Atryn became the first therapeutic protein produced by a transgenic animal to be approved by the FDA. Atryn is an antithrombin expressed in the milk of goats, and was designed to be used in therapy to prevent blood clotting.

Genetically engineered therapeutic antibodies with improved therapeutic potential

One class of recombinant protein has notably been put to therapeutic use: genetically engineered antibodies. As detailed in Section 4.4, each of us has a huge repertoire of different antibodies that act as a defense system against innumerable foreign antigens. Antibody molecules function as adaptors: they have binding sites for foreign antigens at the variable end, and binding sites for effector molecules at the constant end. Binding of an antibody may be sufficient to neutralize some toxins and viruses; more usually, the bound antibody triggers the complement system and cell-mediated killing.

Artificially produced therapeutic antibodies are designed to be mono-specific (specific for a single antigen). Traditional monoclonal antibodies (mAbs) are secreted by *hybridomas,* immortalized cells produced by fusing antibody-producing B lymphocytes from an immunized mouse or rat with cells from an immortal mouse B-lymphocyte tumor. Hybridomas are propagated as individual clones, each of which can provide a permanent and stable source of a single mAb.

The therapeutic potential of mAbs produced like this is, unfortunately, limited. Rodent mAbs, raised against human pathogens for example, have a short half-life in human serum, often causing the recipient to make anti-rodent antibodies. And only some of the different classes can trigger human effector functions.

Genetically engineered antibodies

To make rodent monoclonal more stable in humans, some or all of the rodent protein sequence is replaced by the human equivalent. That happens by genetic engineering at the DNA level: coding DNA sequences encoding part or all of the rodent antibody are replaced by the equivalent human sequences, and the altered coding DNA is expressed to make the desired antibody. The first such attempts were designed to generate a chimeric V/C antibody containing the original rodent variable chains but human constant regions (Figure 9.14).

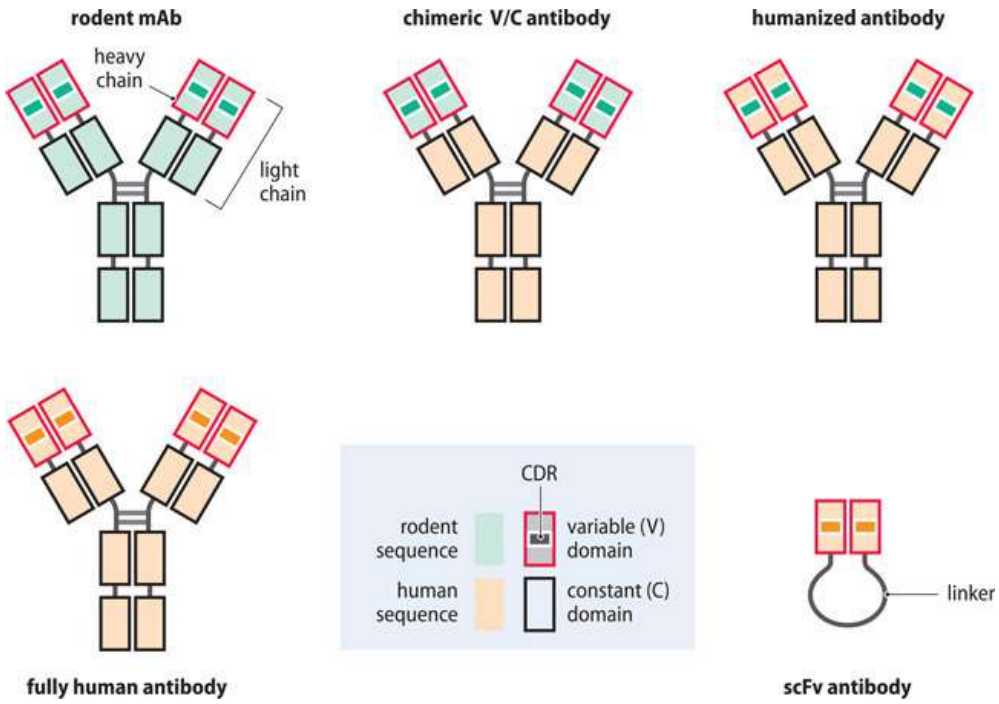


Figure 9.14 Using genetic engineering to make improved therapeutic antibodies. Classical antibodies consist of heavy and light chains with variable (V) and constant (C) domains. Rodent monoclonal antibodies (mAbs) are monospecific antibodies synthesized by hybridomas. Chimeric V/C antibodies are engineered to have human constant domains joined to rodent variable domain sequences (containing the critically important hypervariable complementarity-determining region, CDR). In humanized antibodies all the sequence is human except the hypervariable CDR. Fully human antibodies are obtained by different routes (see text). Single-chain antibodies have also been made, with two variable domains only, connected by a linker peptide. These single-chain variable fragment (scFv) antibodies are particularly well suited to working within the reducing environment of cells; they can serve as ***intrabodies*** (intracellular antibodies) by binding to specific antigens within cells. Depending on the length of the linker, they bind their target as monomers, dimers, or trimers. Multimers bind their target more strongly than monomers.

Subsequently, *humanized antibodies* were constructed in which all the rodent sequence was replaced by human sequence, except for the complementarity-determining regions (CDRs), the hypervariable sequences of the antigen-binding site (see Figure 9.14). More recently, it has been possible to prepare *fully human antibodies* by different routes. For example, mice have been genetically manipulated to delete their immunoglobulin loci and replace them with an

artificial chromosome containing the entire human heavy-chain and g light-chain loci so that they can make fully human antibodies only.

From inauspicious beginnings in the 1980s, mAbs have become the most successful biotech drugs ever, being used to treat a variety of common genetic diseases (see **Table 9.6**). The market for mAbs is the fastest-growing component of the pharmaceutical industry, and of the therapeutic mAbs currently in use, the eight bestsellers are expected to generate together an annual income of more than US $170 billion by the end of 2021. Many additional mAb products are in the pipeline. Of the FDA-approved mAbs, most are partly or fully human and the majority are aimed at treating autoimmune or immunological disease or cancers. In the latter case, the latest antibodies are being developed as antibody–drug conjugates so that the antibodies deliver powerful toxins to kill cancer cells.

Intracellular antibodies (intrabodies)

Simplified antibodies containing just the variable sequences important in antigen recognition can function inside cells and bind to specific intracellular antigens. They are produced by genetic engineering: the appropriate genetic construct is made, then transfected into suitable cells to produce the desired intrabody. They therefore complement standard-size antibodies (which bind to epitopes on cell surfaces), and their therapeutic potential is being tested. There are two significant categories, as listed below.

- *Single-chain antibodies*. Engineered to have a one variable chain, single-chain variable fragment (scFv) antibodies have almost all the binding specificity of a mAb (see Figure 9.14). They can be made on a large scale in bacterial, yeast or even plant cells. Unlike multichain standard antibodies, scFv antibodies are stable in the reducing environment within cells, and well suited to acting as intrabodies. They are designed to bind to specific target molecules within cells. As required, they can be directed towards specific subcellular compartments.
- *Nanobodies* (*single-domain antibody fragments*). The starting point was the discovery that camelids (camels, llamas and so on) have fully functional antibodies that lack light chains, and their heavy chain-only antibodies have a single variable domain. Thereafter, cloned, isolated single variable domains were found to have full antigen-binding capacity and to be very stable compared to normal antibodies.

Intrabodies can carry effector molecules that perform specific functions when antigen binding occurs. However, for many therapeutic purposes they are designed simply to block specific protein–protein associations within cells. As such, they complement conventional drugs. Protein–protein interactions usually occur across large, flat surfaces and are often unsuitable targets for small molecule drugs (that normally operate by fitting snugly into clefts on the surface of macromolecules). Potentially promising therapeutic target proteins for

intrabodies include mutant proteins that tend to misfold in a way that causes neurons to die, as in various neurodegenerative diseases including Alzheimer, Huntington, and prion diseases.

9.3 PRINCIPLES OF GENE AND CELL THERAPY

Gene therapy involves the direct genetic modification of cells to achieve a therapeutic goal. The genetic modification can involve the insertion of DNA, RNA, or oligonucleotides. Gene therapy can be classified into two types, according to whether somatic cells or germline cells are genetically modified. Somatic cell gene therapy seeks to modify specific cells or tissues of the patient in a way that is confined to that patient. Germline gene therapy would produce a permanent modification that can be transmitted to descendants; this could be achieved by modifying the DNA of a gamete, zygote, or early embryo.

Germline gene therapy that involves modifying nuclear DNA is widely banned in humans for ethical reasons, but, as described in Section 9.4, replacement of mutant mtDNA by normal mtDNA in oocytes or zygotes is licensed in some countries to prevent transmission of certain mitochondrial DNA disorders. Ethical issues relating to gene therapy are discussed in Section 11.5.

Gene therapy has had a checkered history. Tremendous initial excitement—and quite a bit of hype—was followed by a fallow period of disappointing results and safety concerns (with unexpected deaths of patients arising from unforeseen deficiencies in the treatment methods). More recently, there has been a greater appreciation of safety risks, and significant successes.

In this section and Section 9.4, we are mostly concerned with gene therapy for inherited disorders, which has focused predominantly on recessive Mendelian disorders. Cancer gene therapy and other approaches to treating cancer are described in Chapter 10.

The first real successes of gene therapy were not achieved until the early 2000s and involved treating very rare cases of severe combined immunodeficiencies. They took advantage of previous experience of bone marrow transplantation, which is effectively a crude type of hematopoietic stem cell therapy.

As we describe later, *stem cells* are cells that have the property of being able to renew themselves and also being able to give rise to more specialized cells. For many types of gene therapy, it is important to maximize gene transfer into appropriate stem cells in the patient. Cell therapies based on the genetic modification of stem cells are also fundamental in *regenerative medicine, where* the object is to treat disease by replacing cells or tissues that have been lost through disease or injury.

In this section we consider the principles underlying gene and cell therapy. In Section 9.4 we deal with the progress made, and discuss future prospects.

Two broad strategies in somatic gene therapy

The cells targeted in somatic cell gene therapy are normally those directly involved in the pathogenic process, but in many cancer gene therapy trials the object has been to genetically

modify *normal* cells in a patient to provoke specific immune responses and killing of harmful cells.

Using molecular genetic approaches to treat disease might involve many different strategies. But at the level of the diseased cells there are two basic strategies: disease cells are simply genetically modified in some way so as to alleviate disease, or they are selectively killed. Within each of the two main strategies are different substrategies, as described below.

Modifying disease cells (Figure 9.15A). According to the molecular pathology, different strategies are used. If the problem is loss of function, a simple solution (in theory) is to add functioning copies of the relevant gene. In genetic disorders in which the pathogenesis results from a gain of function, there is some harmful or toxic gene product within cells. The approach then might be to selectively inhibit the expression of the harmful gene product without affecting the expression of any normal genes. This can often be done by selectively blocking transcription of the harmful mutant gene or by targeting transcripts of the gene so that they are destroyed *(gene silencing)*. Yet other approaches seek to repair a genetic lesion by some type of *genome editing* or find a way of minimizing its effect. We detail the approaches below.

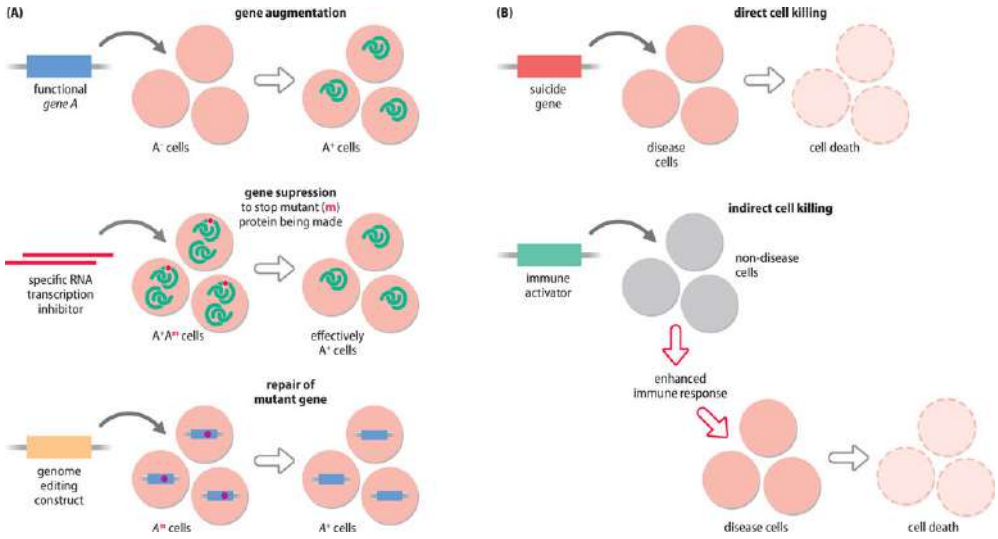


Figure 9.15 Different general types of gene therapy strategy. (A) Therapies aimed at modifying disease cells. *Gene augmentation* therapy (also called gene supplementation or gene addition therapy) can be applied to loss-of-function disorders but is currently limited to treating recessive disorders (in which the disease results from a lack, or an almost complete lack, of some gene product). The object is simply to transfer a cloned working gene copy into the cells of the patient in order to make some gene product that is lacking. *Gene suppression therapy* can be applied to disorders that result from positively harmful gene products. If the disease is caused by a gain-of-function mutation that produces a harmful mutant gene product, A^m, in addition to the normal gene product, A^+, one might try to specifically inhibit the expression of the mutant allele without inhibiting expression of the normal allele. The same approach can be applied to treating autoimmune and infectious diseases. *Genome editing* can be used to repair a DNA lesion, converting the sequence of the mutant allele, A^m, to a normal allele sequence, A. (B) Therapies aimed at killing harmful cells. The overwhelming application has been in cancer

gene therapy trials, either seeking to kill cancer cells directly (by inserting and expressing cloned genes that give rise to some cytotoxic product and cell death) or indirectly (by transferring genes into non-disease cells, such as immune system cells, to provoke an immune response directed at tumors).

Killing disease cells (Figure 9.15B). This approach has frequently been used in cancer gene therapy trials. Traditional cancer treatments have often relied on killing disease cells by using blunt instruments, such as high-energy radiation and harmful chemicals that selectively kill dividing cells. Gene therapy approaches can kill harmful cells either directly or by modifying immune system cells to enhance immune responses that can kill the harmful cells.

The delivery problem: designing optimal and safe strategies for getting genetic constructs into the cells of patients

In gene therapy, a therapeutic *genetic construct* of some type—often a cloned gene, but sometimes RNA or oligonucleotides—is transferred into the cells of a patient. (A nucleic acid molecule introduced in this way is often referred to as a **transgene**.) Depending on the disease, the type of cells to be targeted can be very different, and different strategies are needed. And, depending on the target cells, some disorders are easier to treat in principle than others.

Consider access to the desired target cells. Some cells and tissues—notably blood, skin, muscle, and eyes—are very accessible; others, such as brain cells, are not easily accessed. Then there is the question of overcoming various barriers that impede the transfer and expression of genetic constructs. Strong immune responses constitute important barriers. And, as we will see, mechanical barriers can also be important.

For target cells, another significant factor is the extent to which the cells divide. Regular cell division is required to replenish short-lived cells, such as blood and skin cells, unlike in the case of long-lived cells, such as terminally differentiated muscle cells. The distinction is important. For nondividing cells the key parameters would simply be the efficiency of transfer of the therapeutic construct into the cells of the patient and the degree to which the introduced construct was able to function in the expected way. However, for dividing cells we also need to take into account what happens to the descendant cells.

Even if we were to achieve significant success in getting the desired genetic construct into short-lived cells, the cells that have taken up the genetic construct are going to die and will be replaced by new cells. A small minority of the cells, **stem cells**, divide to continuously replenish cells lost through aging, illness, or injury (see **Box 9.1** for a brief overview of stem cells). To ensure that copies of the therapeutic construct keep getting into newly dividing cells, therefore, it would be best to target the relevant stem cells if possible, and get the therapeutic construct integrated into chromosomes (so that it gets replicated, allowing copies to be passed to both daughter cells at cell division).

BOX 9.1 AN OVERVIEW OF STEM CELLS AND ARTIFICIAL EPIGENETIC REPROGRAMMING OF CELLS

Stem cells have two essential properties: they can self-renew and they can also give rise to more **differentiated** (more specialized) cells. As well as undergoing normal (symmetric) cell divisions, stem cells can undergo asymmetric cell division to give two *different* daughter cells. One daughter cell is identical to the parent stem cell, allowing self-renewal; the other daughter cell is more specialized and can undergo further rounds of differentiation to give terminally differentiated cells (Figure 1).

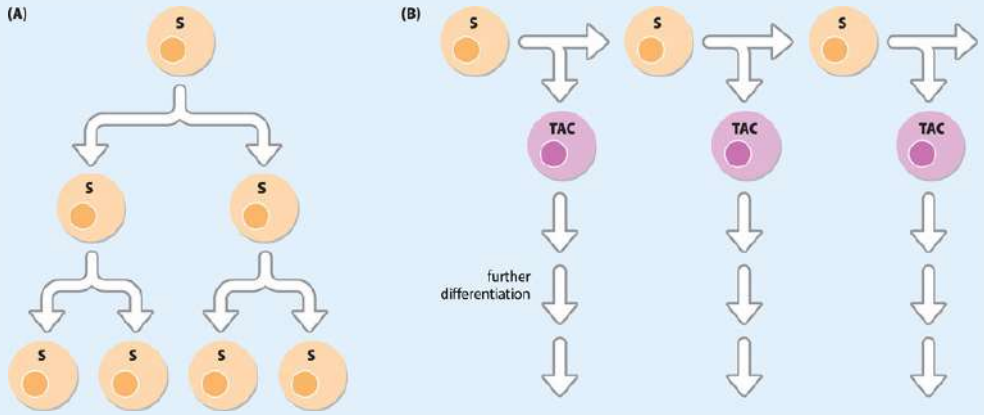


Figure 1 Symmetric and asymmetric stem cell divisions. (A) Populations of stem cells (S) can expand quickly by symmetrical cell division during growth and when there is a rapid need for new stem cells (to replace cells lost through disease or injury). (B) Stem cells give rise to more differentiated cells by asymmetric cell divisions—the stem cell produces two different daughter cells. One daughter cell is a stem cell identical to the parent cell. The other is a stem cell derivative, sometimes called a **transit amplifying cell** (TAC), that is more differentiated than its sister or parent cell and can subsequently undergo additional differentiation steps to form a terminally differentiated cell.

Different classes of stem cell are used in experimental investigations. Some are cultured *somatic* stem cells derived from naturally occurring somatic stem cells in the body. They can give rise to a limited number of differentiated cell types. The other major class are artificially created *pluripotent* stem cells that have the capacity to give rise to all of the different cell types in the body. Two major types of cultured pluripotent stem cells are described below.

- **Embryonic stem cell (ESC) lines** are artificially derived from naturally pluripotent cells of the very early embryo (which, however, are not stem cells, being *transient* cells that will change into more specialized cells during development).
- **Induced pluripotent stem cells (iPSCs)** are obtained by artificially changing the normal epigenetic settings of easily obtained cells such as skin cells; this is a type of artificial epigenetic reprogramming.

SOMATIC STEM CELLS

These cells—also called adult stem cells or tissue stem cells—occur naturally in the body, and help replace cells with naturally short life spans (notably in the blood, skin, intestines, and testis) or help replenish cells lost in disease or injury. (Our powers of tissue regeneration are, however, rather limited.)

Most somatic stem cells give rise to a limited set of differentiated cells. Some, such as spermatogonial stem cells, are unipotent, giving rise to a single type of differentiated cell. Others are multipotent, being able to give rise to several different classes of differentiated cells. For example, hematopoietic stem cells can give rise to all types of blood cell as well as to certain tissue cell types (Figure 9.17). Cultured somatic stem cell lines have been used for studying differentiation, and purified populations of genetically modified somatic stem cells, notably hematopoietic stem cells, have been used in gene therapy.

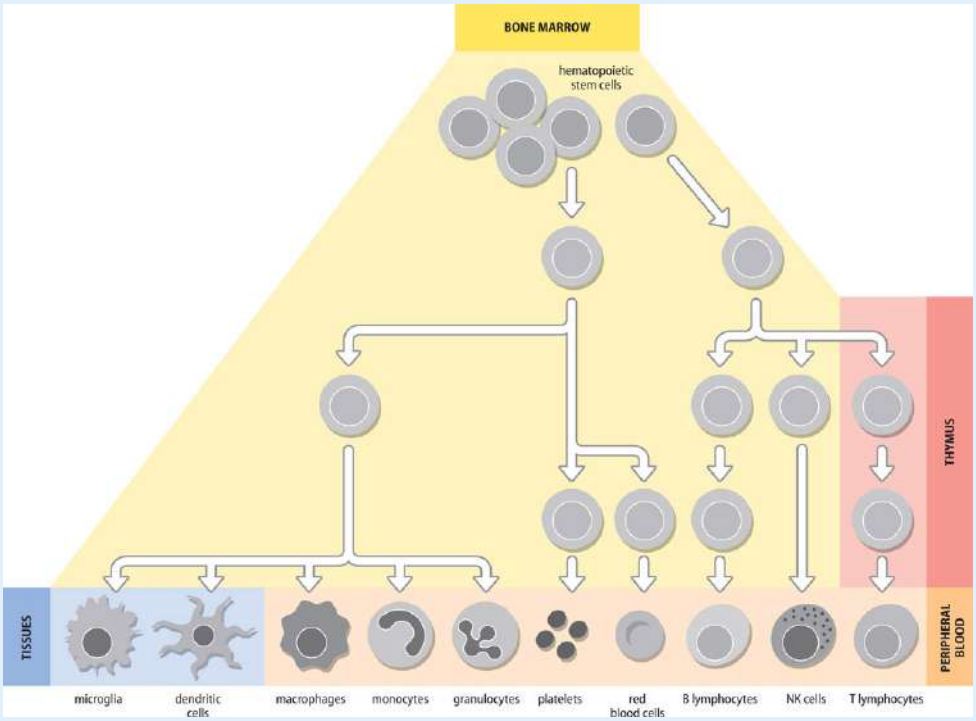


Figure 9.17 All blood cells and some tissue immune system cells originate from hematopoietic stem cells. All differentiated blood cells have limited life spans, and there is a continuous cycle of cell death and cell replacement. The replacement blood cells are derived from hematopoietic stem cells that are particularly concentrated in the bone marrow. Hematopoietic stem cells also give rise to some tissue cells, including tissue macrophages (such as microglia, the resident macrophages of the brain and spinal cord) and dendritic cells (a class of immune system cells that work in antigen presentation in varied tissues). NK cells, natural killer cells.

EMBRYONIC STEM CELL LINES

The mammalian zygote and cells descending from it through the first few cleavage divisions) are entirely unspecialized and are said to be *totipotent*—they can give rise to every type of cell in both the embryo and in the extra-embryonic membranes. Subsequently, the *blastocyst* forms as a hollow ball of cells with two quite distinct layers: an outer layer of cells known as the *trophoblast* (which will give rise to the extraembryonic membranes such as the chorion and the amnion), and a group of inner cells, the *inner cell mass*, located at one end of the blastocyst (**Figure 2**).

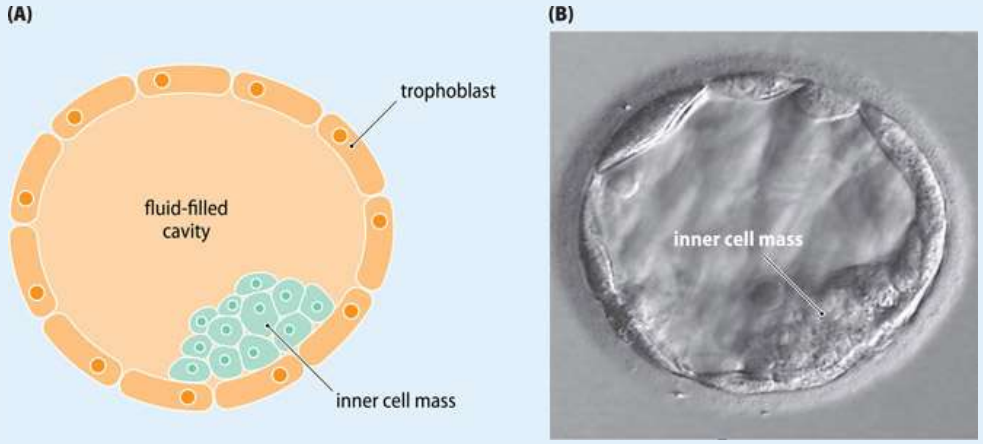


Figure 2 Blastocyst structure and the inner cell mass. (A) The blastocyst is a hollow, fluid-filled ball of cells with two distinct cell populations: an outer cell layer (the trophoblast) and an inner group of cells located at one end (the inner cell mass). (B) A 6-day-old human blastocyst containing about 100 cells, showing the location of the inner cell mass. (B, Courtesy of M. Herbert, Newcastle University, UK.)

Cells from the transient inner cell mass are pluripotent and can be cultured to make a pluripotent embryonic stem cell (ESC) line. ESCs can then be experimentally induced to make desired types of differentiated cell (including derivatives of the germ cell layers—ectoderm, mesoderm, and endoderm—and also germ cells). They have been vitally important for making animal (mostly, mouse) models of disease, as described in Box 9.2. Human ESC lines are produced from cells derived from surplus embryos in assisted reproduction *(in vitro* fertilization; IVF) clinics. They may have promise in cell therapy if immune responses in recipients are minimized in some way, but because human ESC lines are derived from human embryos, the creation of new ESC lines remains controversial.

INDUCED PLURIPOTENT STEM CELLS AND CELL REPROGRAMMING

For decades, cell differentiation in mammals was thought to be irreversible. Then a cloned sheep called Dolly proved that terminally differentiated mammalian cells could be reprogrammed to become unspecialized cells resembling the pluripotent cells of the early embryo. Cloning mammals is, however, extremely arduous and technically difficult.

Alternative, comparatively simple, methods can be used to re-set the epigenetic marks of a cell. For example, by providing certain key transcription factors, or by inducing the cells to

make them, terminally differentiated mammalian cells may be induced to *dedifferentiate* (to give less specialized cells), or to *transdifferentiate* (to give specialized cells of a different type)

Like ESCs, induced pluripotent stem cells (iPSCs) can be directed to differentiate into more specialized cells (Figure 3). Because iPSCs may retain some characteristics of their progenitor cells, they are less robust than ESCs.

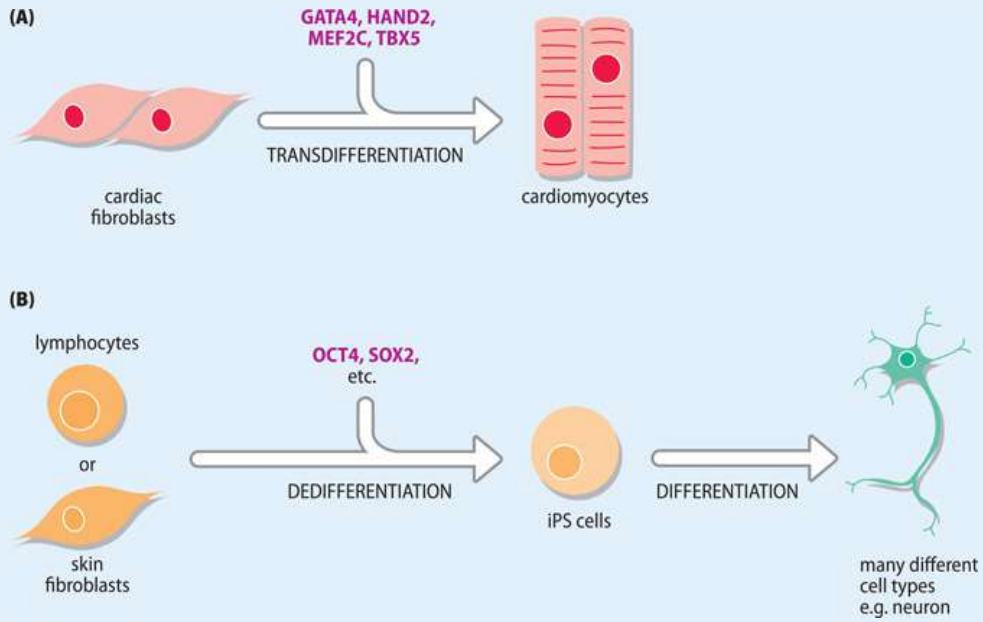


Figure 3 Cell reprogramming. (A) An example of direct reprogramming. Here, certain cardiac transcription factors (in pink) induce the transdifferentiation of fibroblasts to myocytes. (B) Reprogramming to give induced pluripotent stem cells (iPSCs) with the use of transcription factors, such as OCT4 and SOX2, that are important in embryonic development and pluripotency. By providing suitable transcription factors, the resulting iPSCs can be induced to differentiate to give a desired cell type.

From a medical perspective, iPSCs offer two exciting applications: human cellular models of disease and genetically modified cells for therapeutic purposes. Animal disease models have been very valuable, enabling the use of invasive studies to infer the molecular basis of human disease, and p[1]roviding frontline testing of new therapies. But they are only *models;* they quite often show important differences from humans. Accessible skin cells from a patient can now be reprogrammed to become iPSCs that can then be directed to differentiate into cells relevant to the disease process (such as normally inaccessible neurons for a neurodegenerative disorder). Genetically impaired disease cell lines are miniature disease models, useful for drug screening (testing for toxicity, efficacy, and so on) and for studying the molecular basis of disease in *human* cells.

In providing genetically modified cells for therapeutic purposes, iPSCs have the advantage that they can be made from the cells of a patient, genetically modified, and then returned to the patient *without provoking an immune response*. Successful artificially

induced reprogramming of human cells may transform the prospects of using dedifferentiated human cells therapeutically. We describe this aspect in more detail in section 9.4.

Efficiency and safety aspects

In any gene delivery system used in gene therapy, two key parameters are fundamental: efficiency and safety. Most gene therapy methods rely on transferring genes into the cells of a patient and expressing them to make some product. For the gene delivery method to be effective, it is important to maximize transfection efficiencies for the optimal target cells and to get long-lasting high-level expression of the therapeutic genes.

For disorders where target cells are short-lived, targeting the relevant stem cells should maximize the chances of success. However, stem cells occur *in vivo* at very low frequencies. For blood disorders, happily, it is possible to obtain preparations of bone marrow cells or peripheral blood lymphocytes from patients, grow the cells in culture, and enrich for hematopoietic stem cells. The purified cells can be genetically modified in culture to overcome a genetic defect and then returned to the patient, a type of *ex vivo* gene therapy, as described in the next section.

As we describe below, viral vectors are commonly used to get therapeutic gene constructs into cells at high efficiency, and they often allow high-level expression of the therapeutic transgenes. Some viral vectors are deliberately used because they are adept at getting DNA inserted into chromosomes, which is important when targeting tissues in which cells are short-lived. But the features that make the gene therapy process efficient come with significant safety risks.

One important risk concerns the integration of some therapeutic recombinant viruses into chromosomes—there has often been little control over where they will insert into the genomic DNA of patient cells. They might insert by accident into an endogenous gene and block its function, but the greatest danger is the accidental activation of an oncogene, causing tumor formation.

An additional important risk is that the patient might mount a strong immune or inflammatory response to high levels of what might appear to be foreign molecules. Components of viral vectors might pose such risks, but even if a perfectly normal therapeutic human gene were inserted and expressed to give a desired protein that the patient completely lacked (through constitutional homozygous gene deletion, for example), an immune response might occur if the protein had never been produced by the patient. We enlarge on these issues in Section 9.4.

Different ways of delivering therapeutic genetic constructs, and the advantages of *ex vivo* gene therapy

In gene therapy, a genetic construct is inserted into the cells of a patient by using either viral delivery systems or nonviral methods. Viral vector systems are generally much more efficient than nonviral methods, but they pose greater safety risks.

Using viruses to transfer DNA into human or other animal cells (a process called **transduction**) might be expected to be efficient: over long evolutionary timescales, viruses have mastered the process of infecting cells, and getting their genes to be expressed, often after inserting their genomes into host cell DNA. Depending on the type of virus, a virus may have a DNA or RNA genome that is single-stranded or double-stranded. To be useful for ferrying genes into cells, a virus vector is used, a modified double-stranded DNA copy of the viral DNA or RNA genome (making it easy for a therapeutic DNA to be joined to it to form a recombinant DNA).

Viral vectors for use in gene therapy have been engineered to lack most, and quite often all, of the coding capacity of the original viral genome. The idea is that the recombinant DNA (virus vector plus therapeutic DNA) can nevertheless get packaged into a viral protein coat to make a recombinant virus that is still efficient at infecting cells. In some cases the recombinant viral DNA can integrate into the nuclear genome of a cell, permitting long-lasting therapeutic gene expression; but the integration of vectors poses safety risks. When using other types of virus vector, the recombinant viral DNA does not integrate into the host cell genome; instead, it remains as an extrachromosomal **episome** in cells.

The nonviral transfer of DNA, RNA, or oligonucleotides into human or other animal cells (**transfection**) is much less efficient than viral transduction; the overall amount of transgene expression is therefore more limited. The transfection procedures also do not result in appreciable integration of DNA into the genome of the cell. As a result, transfection has the advantage of greater safety in therapeutic applications, but with reduced efficiency. In addition, transfection methods do not have the same size constraints for the packaged nucleic acid that applies to virus vectors—they can be used to ferry very large nucleic acids.

In vivo and *ex vivo* gene therapy

Some types of gene therapy procedure occur *in vivo:* the transfer of the therapeutic constructs is carried out *in situ* within the patient. Often the therapeutic construct is injected directly into an organ (such as muscle, eye, or brain). It may in some cases be introduced indirectly to target cells. For example, coding sequences of genes important in vision have been successfully delivered into the eyes of patients with hereditary loss of vision with quite good outcomes. We describe below how certain viruses are adept at infecting human cells of a particular type; that property has also been exploited to increase the efficiency of delivering therapeutic genes to the desired target cells.

Because there is no way of selecting and amplifying cells that have taken up (and, in some cases, expressed) the genetic construct, the success of *in vivo* gene therapy is crucially

dependent on the general efficiency of gene transfer and, where appropriate, expression in the correct tissue.

Ex vivo gene therapy means removing cells from a patient, culturing and genetically modifying them *in vitro,* and then returning suitably modified cells back to the patient (**Figure 9.16**). Because the cells of the patient are genetically modified in the laboratory, they have the enormous advantage that the cells can be analyzed at length to identify those in which the intended genetic modification has been successful. The correctly modified cells can then be amplified in culture and injected back into the patient.

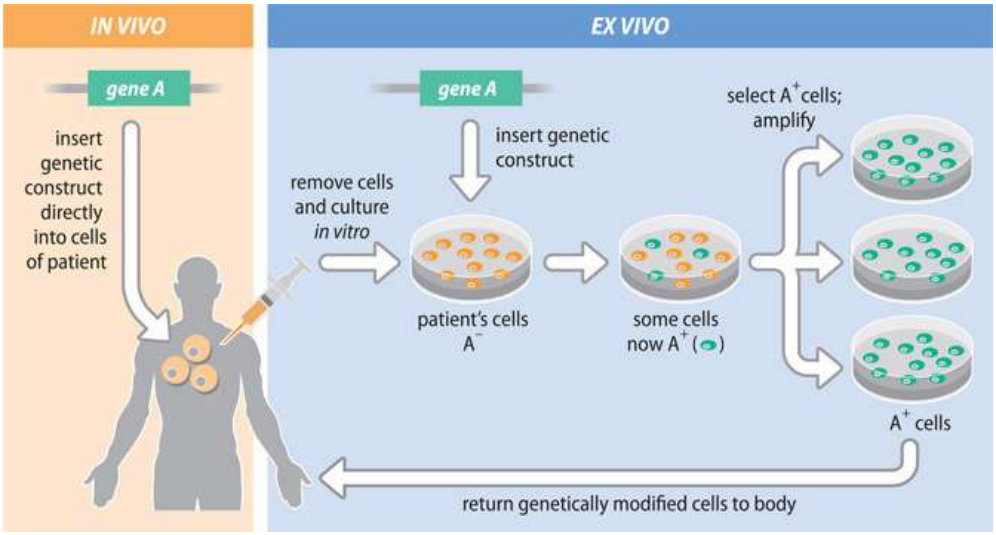


Figure 9.16 ***Ex vivo*** **and** ***in vivo*** **gene therapy**. In *ex vivo* gene therapy, cells are removed from the patient and genetically modified in some way in the laboratory (in this case we illustrate a gene supplementation procedure in which a therapeutic transgene, gene *A*, is expressed to make a gene product, A, that is lacking in the cells of the patient). The modified cells are selected, amplified in culture and returned to the patient. The procedure allows detailed checking of genetically modified cells to ensure that they have the correct genetic modification before they are returned to the patient. For many tissues, this is not possible, and the cells must be modified directly within the patient's body (*in vivo* gene therapy).

In practice, *ex vivo* gene therapy has been directed at certain disorders—mostly blood disorders but also some storage disorders—in which the genetically modified cells are bone marrow cells that have been taken from the patient and then treated in such a way so as to enrich for hematopoietic stem cells. As described later, this procedure has been at the core of a series of successful gene therapies.

Nonviral delivery of therapeutic genetic constructs

Interest in nonviral vector delivery systems has mostly been propelled by safety concerns over the use of viral vectors. Nonviral vector systems are certainly safer—they do not integrate into chromosomes and they are not very immunogenic.

The therapeutic gene is typically carried in a plasmid vector, but transport of plasmid DNAs into the nucleus of nondividing cells is normally very inefficient (the plasmid DNA often cannot enter nuclear membrane pores). Various tricks can be used to help get the plasmids into the nucleus (such as compacting the DNA to a small enough size to pass through the nuclear pores). Because the transfected DNA cannot be stably integrated into the chromosomes of the host cell, nonviral methods of therapeutic gene delivery are more suited to delivery into tissues such as muscle, which do not regularly proliferate, and in which the injected DNA may continue to be expressed for several months.

Different delivery systems can be used. **Liposomes**, synthetic vesicles that form spontaneously when certain lipids are mixed in aqueous solution, have often been used to enclose the desired DNA construct. The lipid coating allows the DNA to survive *in vivo,* bind to cells, and be endocytosed into the cells. Another method uses compacted DNA nanoparticles. Because of its phosphate groups, DNA is a polyanion. Polycations bind strongly to DNA and so cause the DNA to be significantly compacted. Because of their much reduced size, compacted DNA nanoparticles are comparatively efficient at transferring genes to dividing and nondividing cells and have a plasmid capacity of at least 20 kb.

Currently, there is still some way to go for nonviral delivery systems. The efficiency of getting genetic constructs into cells using these methods remains low, as does the expression levels of transfected therapeutic genes. Interested readers can find a recent review at PMID 32580326.

Viral delivery of therapeutic gene constructs: relatively high efficiency but safety concerns

Before we detail how viruses can be used as vectors in gene therapy, let us first consider some properties of viruses. Viruses have a DNA or RNA genome packaged within a protein shell (known as a *capsid*), and in some viruses, called *enveloped viruses*, the protein capsid is in turn enclosed by a lipid bilayer containing viral proteins. Enveloped viruses enter cells either by fusing with the host plasma membrane to release their genome and capsid proteins into the cytosol, or by first binding to cell surface receptors, and then entering via receptor-mediated endocytosis, fusion-based transfer, or endocytosis-based transfer.

Some viruses infect a broad range of human cell types and are said to have a broad *tropism*. Other viruses have a narrow tropism: they bind to receptors expressed by only a few cell types. Herpes viruses, for example, are tropic for cells of the central nervous system. The natural tropism of viruses may be retained in vectors or genetically modified in some way, so as to target a particular tissue.

For an introduced transgene to be expressed, it needs to be ferried to the nucleus. Some viruses can gain access to the nucleus only after the nuclear envelope has dissolved during mitosis. They are limited to infecting dividing cells. Other viruses have devised ways to transfer their genomes efficiently through nuclear membrane pores, so that both dividing and nondividing cells can be infected.

Some viruses are able to integrate their genome into the genome of the host cell. They include retroviruses, whose genome consists of a single RNA strand, such as lentiviruses and gammaretroviruses. They are able to convert their RNA into a single-stranded cDNA using a viral reverse transcriptase; the single DNA strand is then copied into a double-stranded DNA (replicative form) that can integrate into the host genome using viral enzymes. Other viruses, such as adenoviruses and adeno-associated viruses, do not integrate into the genome.

Virus vectors used in gene therapy

TABLE 9.7 FOUR MAJOR CLASSES OF VIRAL VECTORS THAT HAVE BEEN USED IN GENETHERAPY PROTOCOLS

	Virus class	Viral genome	Cloning capacity	Target cells (D or ND cells)	Transgene expression	Vector yield*; other comments
INTEGRATING	Gamma-retroviruses	ssRNA;˜8-10kb	7-8 kb	D cells only	long-lasting	moderate; risk of oncogene activation
	Lentiviruses (notably HIV)	ssRNA;˜9kb	Up to 8 kb	D and ND cells; tropism varies	long-lasting and high-level	high; risk of oncogene activation
NON-INTEGRATING	Adenoviruses	dsDNA; 38-39 kb	often 7.5 kb; up to 34 kb	Mostly ND cells	transient but high-level	high; imunogenicity can be a major problem
	Adeno-associated viruses (AAVs)	ssDNA; 5 kb	<4.5 kb	Mostly ND cells	high-level in medium/long-term (year)	high; small size capacity; immunogenicity less than for adenovirus

Abbreviations: ss, single-stranded; ds, double-stranded; D, dividing cells; ND, non-dividing cells.

* High vector yield, 10^{12} transducing units/ml; moderate vector yield, 10^{10} transducing units/ml.

Four major classes of virus have been used as vectors for gene therapy (see **Table 9.7** for a summary of their properties). The big advantage of using virus vectors is the efficiency in getting transgenes into cells, which far exceeds that of nonviral transfection methods. Over evolutionary time scales, viruses have become adept at infecting cells, and in expressing their

genes within cells. There can be significant safety concerns, however, when using viruses as compared with non-viral transfer methods, and we describe these later.

Vectors based on integrating viruses allow therapeutic genes to be inserted into chromosomes of cells and to be passed on to any descendant cells (an important advantage if the target cells are blood cells or other cell types that have a high rate of cell turnover). For both gammaretroviruses and lentiviruses, the vector is made by isolating viral replicative forms that consist of double-stranded DNA, and genetically modifying them in various ways.

Non-integrating vectors are traditionally based on DNA viruses, and they can be especially useful when the object is to get high-level expression in nondividing target cells, such as muscle. Vectors based on adenovirus have been popular because they can permit very high levels of gene expression, but here, too, there have been safety issues (which relate to their immunogenicity). Safer vectors based on adeno-associated virus (AAV) have subsequently been widely used. We detail the use of viral vectors in gene therapy and the issue of safety concerns in Section 9.4.

The importance of disease models for testing potential therapies in humans

Cellular disease models can be very helpful in understanding the molecular basis of disease, and can be important in drug screening and drug toxicity assays. Recent advances in stem cell technology have allowed the production of a wide range of human cellular disease models. Readily accessible blood or skin cells from a patient can be genetically reprogrammed so that they are converted to some desired cell type that is principally involved in the pathology, such as normally inaccessible neurons. We cover the relevant technology—*induced pluripotent stem cells*—in Section 9.4.

To test novel therapeutic approaches, a robust whole-animal model of disease is necessary. Some animal models of disease, such as the *mdx* mouse model of muscular dystrophy, originated by spontaneous mutation, but the vast majority are artificially generated by genetic manipulation. Primate models might be expected to be the most faithful disease models, but for decades the preferred disease models have been rodent models, notably mice. There are several reasons: rodents breed quickly and prolifically; they are reasonably closely related to humans, sharing 99 % of our genes; maintaining rodent colonies is not too expensive; and there are fewer ethical concerns than with primate models. An additional compelling reason is that for decades certain important genetic manipulation technologies have effectively been available in mice only.

The vast majority of rodent disease models have been created by genetically modifying the germ line, in which foreign DNA is typically engineered into the chromosomal DNA of germline cells. One way is to make a **transgenic animal** disease model by inserting a transgene (= any foreign DNA) into the zygote. This approach can be used in a wide range of different animals.

A second, powerful, technology relies on first genetically modifying the genome of intact embryonic stem (ES) cells in culture in a precise, pre-determined way. The modified ES cells

are then transferred into the early embryo in order to produce an animal with genetically modified cells, including modified germline cells. Certain mouse ES cell lines have been particularly amenable to this genetic modification (which is why mouse disease models are so prevalent). The technology is so sophisticated that we can, in principle, make any desired change to the genome sequence of a mouse—even substituting a single nucleotide—at essentially any position we choose. See **Box 9.2** for the salient details.

BOX 9.2 TWO POPULAR WAYS OF MAKING MOUSE DISEASE MODELS

TRANSGENESIS THROUGH PRONUCLEAR MICROINJECTION

One important route for making transgenic mice (or other transgenic animals) is to inject a transgene into the zygote so that the exogenous DNA gets into the genome of the zygote. There is usually no control over where the transgene integrates. The resulting animal will have the transgene in all cells and can transmit it to future generations (Figure 1).

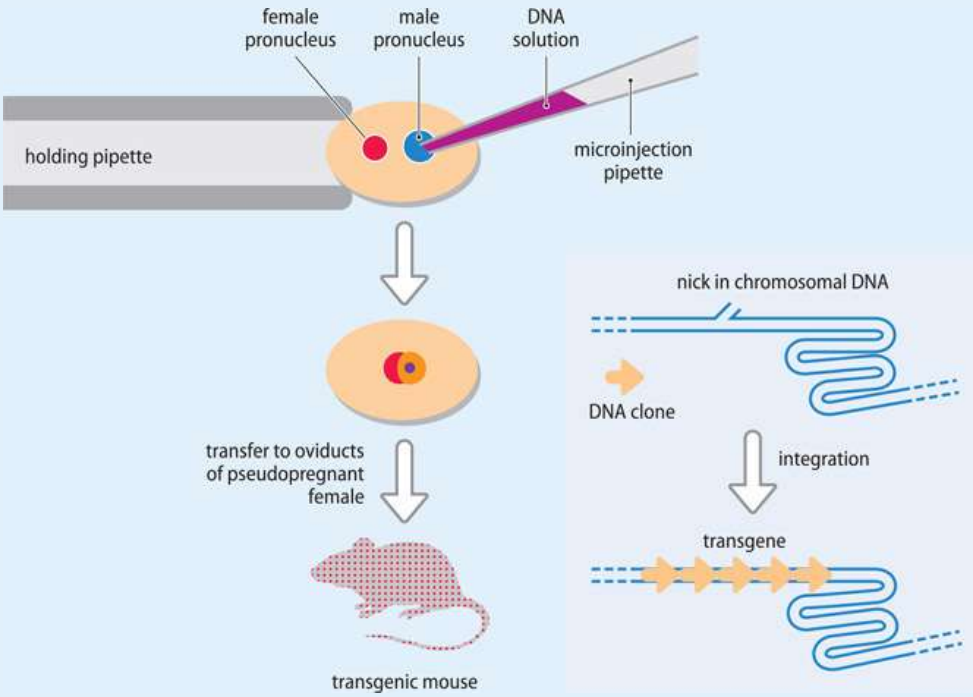


Figure 1 Construction of transgenic mice by pronuclear microinjection. A fine-pointed microinjection pipette is used to pierce first the oocyte and then the male pronucleus (which is bigger than the female pronucleus), delivering an aqueous solution of a desired DNA clone. The introduced DNA integrates at a *nick* (single-stranded DNA break) that has occurred randomly in the chromosomal DNA. The integrated transgene usually consists of multiple copies of the DNA clone. Surviving oocytes are reimplanted into the oviducts of foster females. DNA analysis of tail biopsies from resulting newborn mice checks for the presence of the desired DNA sequence.

This method is often used for modeling dominantly inherited disease due to gain of function or overexpression. In the former case, for example, the trans-gene might often be a mutant human cDNA with an attached promoter sequence to drive expression of the mutant protein in the same cells as those in which it is expressed in humans. Larger transgenes are possible, too, and have sometimes included artificial human chromosomes.

MAKING PRECISE GENETIC MODIFICATIONS IN INTACT EMBRYONIC STEM CELLS

Another popular way of getting foreign DNA into the mouse germ line begins with cultured **embryonic stem (ES) cells** derived from pluripotent cells of the early mouse embryo. ES cells are immortal and can be used to give rise to all cells of the organism, including gametes. A selected mouse ES cell line is genetically modified in culture and then transferred into an isolated mouse blastocyst. The genetically modified blastocyst is implanted into a mouse, which is bred to obtain genetically modified mice.

Genetic modification of an ES cell line in culture has the big advantage that a very precise change—sometimes just a specific single nucleotide change—can be made to order within any individual gene or locus of interest in intact ES cells. This procedure, known as **genome editing**, requires double-strand DNA breaks to be made at the locus of interest, and two major methods have been used as listed below.

- **Gene targeting**. This is a type of homologous recombination. The specificity is provided by transfecting a *homologous* DNA sequence containing an artificially created desired genetic modification. The normal sequence at the locus of interest is replaced by the introduced sequence with the genetic modification using endogenous nucleases in the cell. An appropriate double crossover will suffice as shown in Figure 2C.

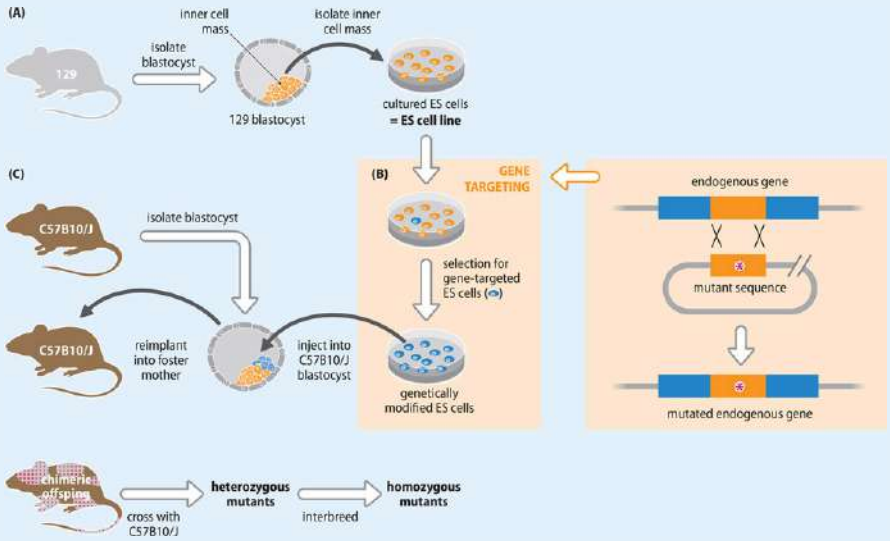


Figure 2 Gene targeting in embryonic stem cells to introduce mutations into the mouse germ line. (A) Embryonic stem (ES) cell lines are made by excising blastocysts from the oviducts of a suitable mouse strain. Cells from the inner cell mass are cultured to eventually give an ES cell line. (B) An ES cell line can be genetically modified in culture by transfecting a linearized plasmid containing a DNA sequence (orange box) that is identical to part of the endogenous target gene, except for a genetically engineered desired mutation (magenta asterisk). Double recombination (X) allows the desired mutation to be introduced into the endogenous gene. (C) The genetically modified ES cells are injected into an isolated blastocyst from another mouse strain with a different coat color, and the blastocyst containing modified ES cells is implanted into a foster mother of the same strain. Subsequent development of the introduced blastocyst can generate chimeric offspring that can be readily identified because they have differently colored coat patches. Backcrossing of chimeras can produce heterozygous mutants (if the genetically modified ES cells have contributed to the germ line); subsequent interbreeding generates homozygous mutants.

- **CRISPR-Cas**. This new method is simple and can be conducted comparatively quickly. The specificity is provided by RNA sequences designed to bind to selected target sequences on opposing DNA strands at the desired locus. Cleavage of the DNA at the desired locus is carried out on both strands using engineered Cas nucleases introduced into the ES cells. We describe this method later in the text.

The techniques for germline modification have been widely used to make loss-of-function mutations to inactivate a gene (these mutations are known as **gene knockouts**, and mice containing them are called knockout mice). Homozygous loss-of-function mouse mutants are often used to model human recessive disorders, but the method also delivers heterozygous mutants that might show pheno-types too. If the homozygous condition is lethal, the mutation can be maintained in heterozygotes, and the mutant strain can be stored for decades by freezing cells in liquid nitrogen. Variant methods can be used to make desired subchromosomal duplications and deletions, translocations, and so on (*chromo-some engineering*).

Inevitably (because it is simpler to do so), most of the artificially created disease models are intended to replicate monogenic disorders. Some good disease models have been produced, and they have been very helpful in allowing us to gain insights into the molecular basis of human diseases, and in testing gene therapies and other new treatments.

When rodent disease models can be inadequate

Rodent models are generally extremely valuable, but are limited in some ways. Mice are small and are less well-suited than larger mammals to physiological analyses. Larger animal disease

models including dog, pig, sheep and primate models have been constructed for some disorders.

Because of species differences, rodent models may quite often fail to replicate some aspects of the human phenotype that they were intended to mimic. In some cases, they are simply inadequate to the task. Disorders such as autism, schizophrenia, and Alzheimer disease cannot be fully replicated in mice (which lack the complex cognitive and social abilities of primates). Many neuroactive drugs have shown early promise in mice but failed in human trials.

As a result of these difficulties, and because of the recent emergence of a transformative technology, genome editing using the CRISPR-Cas system, there has been renewed interest in making primate disease models. Because it also offers interesting therapeutic potential we consider the CRIPSR-Cas method of genome editing in the next section.

9.4 GENE THERAPY FOR INHERITED DISORDERS: PRACTICE AND FUTURE DIRECTIONS

Gene therapy has had a roller-coaster ride over three decades; periods of over-optimism would be followed by bouts of excessive pessimism in response to significant setbacks. The first undoubted successes were reported in the early 2000s, and the number of successful reports is beginning to increase significantly.

By 2021, the Gene Therapy Clinical Trials Worldwide Database (available at http://www.abedia.com/wiley/) had listed over 3000 such trials. Close to two-thirds of them have been aimed at treating cancers and have often been of limited clinical value—we consider the general difficulties against the broader background of cancer therapy in Chapter 10. Here we focus on gene therapy for inherited disorders where the approach is to modify the disease cells genetically.

Other gene therapy trials have been split almost equally between monogenic disorders, complex cardiovascular diseases, infectious diseases, and other categories. However, only 3 % of the listed trials are phase III trials, in which the efficacy of the therapy is tested on a large scale. Despite the limited number of trials, monogenic diseases have always been high on the gene therapy agenda, and the first definitive successes have been in that area.

Multiple successes for *ex vivo* gene supplementation therapy targeted at hematopoietic stem cells

Successful *ex vivo* gene therapy trials have been carried out for various blood disorders and some storage disorders by targeting bone marrow cells or peripheral blood lymphocytes enriched for hematopoietic stem cells. Our blood cells are short-lived and need to be replaced by new cells derived from self-renewing hematopoietic stem cells. These cells, which are found mostly in the bone marrow (and to a smaller extent in peripheral blood), give rise to all of the many different types of blood cell and also to some tissue cells with immune functions (**Figure 9.17**).

For some disorders treated in this way, alternative treatments have sometimes been used, but they are either very expensive or very risky (see below). For some blood disorders, treatment with purified gene product (such as recombinant proteins) is an option, but it is extremely expensive. Bone marrow transplantation has occasionally been used.

In **allogeneic** bone marrow transplantation the donor is often a family relative, such as a sibling, but complete HLA matching of donor and patient is rare (even for siblings there is only a 1 in 4 chance) and sometimes transplantation is attempted using partial HLA matching between donor and recipient. That may result in a severe *graft-versus-host disease,* in which immune system cells originating from the donor bone marrow interpret the cells of the patient as being foreign and mount a strong immune response against them. As a result, the procedure carries a 10–15 % mortality risk that increases to 35 % if the recipient has previously received irradiation treatment (in an attempt to kill many of the original hematopoietic stem cells, so that the transplanted stem cells might expand to become the dominant type).

The advantange of *ex vivo* gene therapy here is two-fold. It is significantly less expensive than using purified proteins. Secondly, it is much less risky than bone marrow transplantation because the genetically modified transplanted cells derive originally from the patient (**autologous** transplantation).

Safety issues in gammaretroviral integration

The first gene therapy successes came in treating severe immunodeficiencies. In severe combined immunodeficiency (SCID) the functions of both B and T lymphocytes are defective. Affected individuals have virtually no functioning immune system and are extremely vulnerable to infectious disease.

The most common form of SCID is X-linked; inactivating mutations in the *IL2RG* gene means a lack of the g_c (common gamma) subunit for multiple inter-leukin receptors, including interleukin receptor 2. (Lymphocytes use interleukins as **cytokines** or chemical messengers that help in intercellular signaling, in this case between different types of lymphocyte and other immune system cells; lack of the g_c cytokine receptor subunit has devastating effects on lymphocyte and immune system function.) Another common form of SCID is due to adenosine deaminase (ADA) deficiency; the resulting buildup of toxic purine metabolites kills T cells. B-cell function is also impaired because B cells are normally regulated by certain types of regulatory T cell.

The first SCID gene therapy trials involved *ex vivo* gammaretroviral transfer of *IL2RG* or *ADA* coding sequences into autologous patient cells. To aid the chances of success, bone marrow cells from the patient were further enriched for hematopoietic stem cells by selecting for cells expressing the CD34 surface antigen, a marker of hematopoietic stem cells (**Figure 9.18**). By 2008, 17 out of 20 patients with X-linked SCID and 11 out of 11 patients with ADA-deficient SCID had been successfully treated and retained a functional immune system (for more than nine years after treatment in the earliest patients).

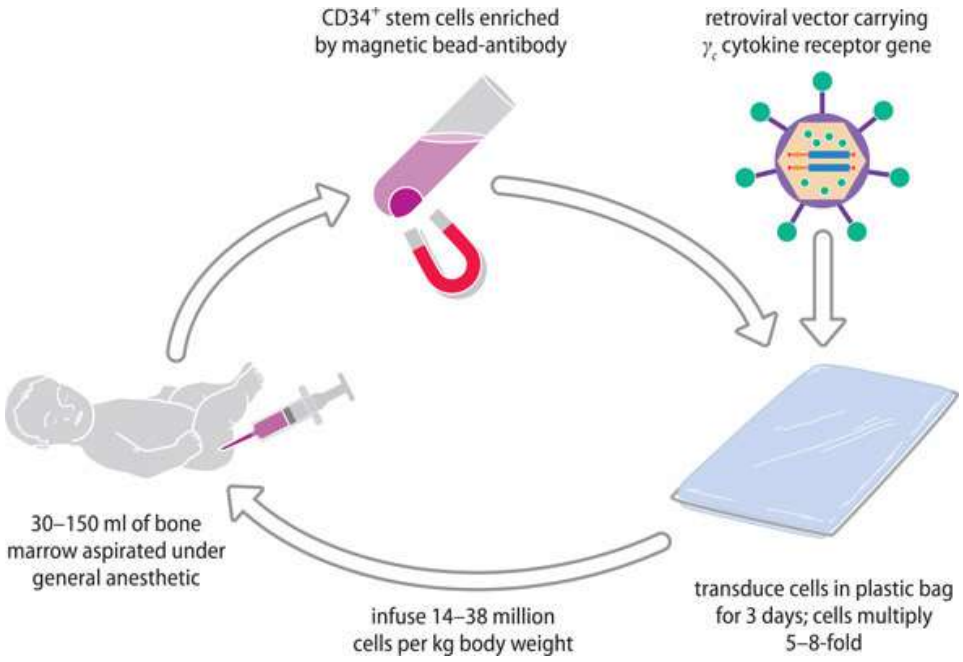


Figure 9.18 The first successful gene therapy: *ex vivo* gene therapy for X-linked severe combined immunodeficiency disease. Bone marrow cells were removed from the patient and antibody affinity was used to enrich for cells expressing the CD34 antigen, a marker of hematopoietic stem cells. To do this, bone marrow cells were mixed with paramagnetic beads coated with a CD34-specific monoclonal antibody; beads containing bound cells were removed with a magnet. The transduced stem cells were expanded in culture before being returned to the patient. Readers interested in the details can find them at PMID 10784449 and 11961146.

Although the use of integrating retrovirus vectors was beneficial in terms of efficiency, the chromosomal insertion of the transgenes was unsafe and led to the development of leukemia in several patients. The same kind of approach has been successfully applied to some other blood disorders. However, oncogene activation in some other cases also led to leukemia in patients or silencing of the inserted transgenes.

It has become clear that gammaretroviral vectors have a pronounced tendency to integrate close to transcriptional start sites, and the long terminal repeats carry very powerful promoter and enhancer sequences that can readily activate the expression of neighboring host cell genes. Retrovirus technology has been made safer by replacing the powerful virus promoter/enhancers by more moderate mammalian control sequences.

More recently, self-inactivating lentivirus vectors have become the first choice for integrating vectors in gene therapy. They have the advantage of long-lasting high-level expression but are much safer than retroviruses: abnormal activation of an endogenous gene is very rarely triggered when lentivirus vectors integrate into chromosomes, mostly because they do not have the same tendency as gammaretroviruses to insert close to transcriptional start sites.

In vivo gene therapy: approaches, barriers, and recent successes

In vivo gene therapy involves the transfer (usually direct) of a genetic construct into post-mitotic disease cells at specific sites in the body (such as muscles, eyes, brain, liver, lung, heart, and joints). Because the intended target cells are nondividing cells, there is no need to insert genes into chromosomes, and so the viral vectors that are used are typically based on non-integrating DNA viruses.

Delivery using adenovirus and adeno-associated virus vectors

Early *in vivo* gene therapy trials often used adenovirus vectors to transfer therapeutic transgenes. These allow high-level expression, and some adenovirus vectors can accept inserts as large as 35 kb (much larger than the vast majority of full-length human cDNA sequences). However, harmful immune and inflammatory responses have sometimes resulted in fatalities. (Because the vectors are non-integrating, and the expression of introduced transgenes is often somewhat transient, short-term and repeated administration would be necessary for sustained expression, but that only exacerbates the immune response).

The safety problems with adenovirus vectors, has prompted a switch to using adeno-associated viruses (AAVs), which are nonpathogenic and are quite unrelated to adenoviruses (their name comes from their natural reliance on simultaneous infection by a helper virus, often an adenovirus). Their most important advantage is that they can permit the robust *in vivo* expression of transgenes in various tissues over several years while exhibiting little immunogenicity and little or no toxicity or inflammation. Multiple different serotypes of AAV have been isolated, and some have a usefully narrow tropism, such as AAV8 (strongly tropic for the liver). There are two downsides. First, a maximum of just 4.5 kb of foreign DNA can be inserted into an AAV vector. Secondly, neutralizing antibodies may be a problem in some people after repeat exposure to the same AAV serovar.

Amenability of disorders to *in vivo* gene therapy

Different disorders may be more or less amenable to *in vivo* gene therapy, largely depending on the efficiency of transgene transfer and expression. That, in turn, partly depends on different types of barrier. Immunological barriers are particularly important when using recombinant virus vectors: as well as posing safety risks, immunological responses can result in transgene silencing (increased host cell cytokine signaling often attenuates the influence of viral promoters).

In addition to immunological barriers, mechanical barriers can also be a major obstacle. Take cystic fibrosis, a disorder that primarily affects the lungs. Gene delivery to the airways using aerosols might seem a very attractive option, given that lung epithelial cells interface directly with the environment. But a combination of immunological and mechanical barriers makes gene therapy a difficult proposition. Lung epithelial cells are locked together by

intercellular *tight junctions,* and large numbers of macrophages are on patrol, readily intercepting and destroying viral vectors. And to top that, there is a natural layer of mucus on the epithelial surface that becomes thicker in individuals with cystic fibrosis, impeding gene transfer.

Some parts of the body are *immunologically privileged sites* in which immune responses to foreign antigen are much weaker than in most other parts of the body (as a result of blood-tissue barriers or a lack of lymphatics, for example). They include the brain and much of the eyes. Additional advantages of the eyes are their accessibility and also their compactness (compare the need for multiple injections at diverse skeletal muscle sites in disorders such as Duchenne muscular dystrophy).

The liver, too, is a quite accessible organ (via direct injection, injection into the hepatic portal vein, or even injection into a peripheral vein); because it has a primary role in biosynthesis, the liver has become a popular target for gene delivery. A wide range of metabolic disorders are caused by defective synthesis of proteins manufactured in the liver (such as blood clotting factors VIII and IX, which are deficient in hemophilia, and many enzymes in inborn errors of metabolism).

Two early examples of successful *in vivo* gene therapy

Hemophilia B (OMIM 306900) is an X-linked recessive disorder caused by a deficiency of blood clotting factor IX. The disorder can be treated by protein therapy (using clotting factor concentrates), but at huge cost. Remarkably, in a study reported by Nathwani et al in the *New England Journal of Medicine* in 2011 (PMID 22149959) a single intravenous injection of a recombinant AAV construct with a factor IX coding sequence could successfully treat patients with hemophilia for more than a year, even though factor IX expression levels were about 10 % or less of the normal values.

In type 2 Leber congenital amaurosis (OMIM 204100), the principal clinical feature—profound loss of vision—usually presents at birth. In the type 2 form, the blindness results from inactivating mutations in both copies of the *RPE65* gene, causing severe retinal degeneration (*RPE65* encodes a retinal pigment epithelium enzyme). Different *in vivo* gene therapy trials have involved injecting a recombinant AAV construct containing a transgene with the *RPE65* coding sequence into the subretinal space, allowing the transduction of retinal pigment epithelial cells. The trials showed the procedure to be both safe and of considerable clinical benefit. In the largest clinical trial, all patients demonstrated increased pupillary response and increased visual field, and a majority of patients demonstrated improved visual acuity.

An overview of RNA and oligonucleotide therapeutics

Popular therapeutic applications for RNA and/or oligonucleotides are summarized in **Figure 9.19**. All of them work by targeting RNA or oligonucleotide sequences to base pair with complementary RNA or DNA sequences at a disease gene locus in order to obtain some therapeutic benefit. They fill a gap that supplementation (augmentation) gene therapy cannot fill. Supplementation gene therapy has undoubtedly been successful in treating certain recessive monogenic disorders (those where the problem is a genetic deficiency). But it is not suited to treating diseases where the mutant gene makes a positively harmful product. To deal with disorders where pathogenesis results from some type of toxic RNA, or a mutant protein with a gain of function or a dominant-negative effect, RNA and oligonucleotide therapeutics offers two major possibilities:

- *gene suppression/silencing* by specific inhibition of, or induced cleavage of, transcripts of the target gene
- *gene repair* (the pathogenic mutation causing a harmful gene product to be produced is repaired by replacing the mutant sequence with a normal one).

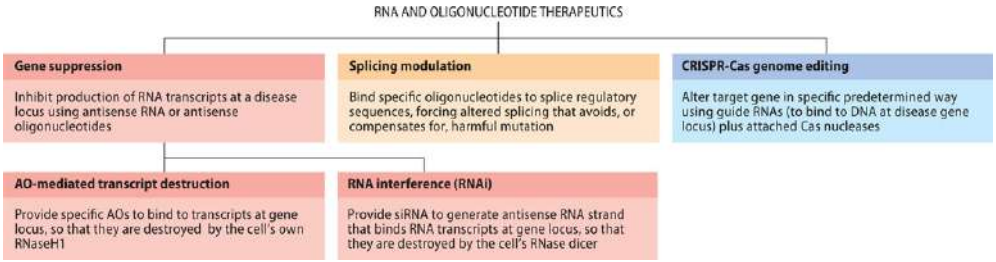


Figure 9.19 An overview of major strategies used in RNA and oligonucleotide therapeutics. All these approaches are focused on RNA transcripts (RNA-targeted therapeutics) except for CRISPR-Cas gene editing, which uses guide RNAs to target DNA sequences. The top right box summarizes the basic CRISPR-Cas method but newer variants called base editing and prime editing use additional enzymes, as described in the text. To make therapeutic oligonucleotides more robust and less susceptible to nuclease attack (when transfected into cells or tissues) their chemical structure is altered, often by having stable phosphorothioate bonds connecting nucleotides instead of phosphodiester bonds and with protective side chains at certain positions on the sugars. siRNA, short interfering RNA, is provided as a RNA duplex but gives rises to a desired antisense RNA within cells. AO, antisense oligonucleotide.

For gene suppression/gene silencing, RNA transcripts from the disease gene locus need to be tagged inside cells by a specific antisense oligonucleotide or RNA in such a way that the bound pathogenic RNA transcripts can be destroyed by a dedicated cellular ribonuclease. It would be optimal, of course, to use a mutant-allele specific antisense oligonucleotide or RNA, but some clinical trials have simply used gene-specific antisense RNA or oligo-nucleotides that bind to transcripts of both the mutant allele and normal allele in affected heterozygotes. The idea here is that sufficient reduction in expression of the mutant allele might be achieved to obtain therapeutic benefit, while because gene suppression is not 100 % efficient, there is

sufficient expression from the normal allele. The two major approaches to gene suppression/gene silencing are listed below.

- *Antisense oligonucleotides (AO)*. The object is to induce ribonuclease RNaseH1 in the cells of an affected individual to selectively destroy transcripts at the disease gene locus. (The natural function of RNaseH1 is to destroy the RNA strands of DNA-RNA hybrids in the cell.) To do this an AO is designed to base pair specifically to transcripts from the disease gene locus, or mutant allele, then transfected into the cells of a patient. The AO must contain a significant number of deoxyribonucleotides so that the RNA-AO hybrid becomes a target for RNaseH1 cleavage of the bound RNA transcripts (often the AO is designed to have a central section of ~10 deoxyribonucleotides flanked by five ribonucleotides on each side).
- *Short interfering RNA (siRNA)*. The object is to exploit a natural innate defense mechanism, RNA interference (RNAi), to specifically destroy RNA transcripts from the disease gene locus or mutant allele. RNA interference is induced after transfecting into the cells of an affected individual a specific siRNA (a double-stranded RNA that will ultimately generate an antisense single-strand RNA) or a gene encoding a precursor of that specific siRNA. The antisense RNA will bind specifically to RNA transcripts from the disease gene locus or mutant allele and thereby tag them to be destroyed by the cells' dicer ribonuclease. In the section following this one we explain RNA interference in detail and show how it has been exploited for therapeutic purposes.

The transformative CRISPR-Cas genome editing method of genome editing (also called **gene editing**) can also be thought of as a type of RNA therapy, even although it works at the DNA level. A potentially powerful therapeutic application of CRISPR-Cas gene editing is in repairing a pathogenic point mutation, replacing the mutant sequence by a normal one. It is crucially dependent on transfecting genes into the desired cells in order to make *guide RNAs* (which are designed to base pair to specific sequences flanking a pathogenic point mutation in the gene of interest) and a Cas (Crispr-associated) nuclease. The object is usually to steer Cas nucleases to cut at pre-determined target sites in the vicinity of the pathogenic mutation within intact cells as a precursor to repairing the mutant gene. We outline this potentially highly promising procedure below.

Another application of oligonucleotide therapeutics, therapeutic splicing modulation, is necessarily limited in scope. The idea is that by designing suitable antisense oligonucleotides to bind to—and thereby blockade—a specific splice junction, an exon with a harmful mutation might be skipped, avoiding the harmful effect of that mutation. Of course, for most mutant genes, this type of exon skipping therapy cannot be applied: even if the induced exon skipping did not induce a translational frameshift, valuable sequence could be expected to be lost. Because of certain unusual characteristics, however, some examples of successful

splicing modulation have been possible in treating Duchenne muscular dystrophy and spinal muscular atrophy (see **Clinical Box 12**).

CLINICAL BOX 12 SPLICE MODULATION THERAPY FOR DUCHENNE MUSCULAR DYSTROPHY AND SPINAL MUSCULAR ATROPHY

Splice modulation therapy has particularly been applied to treating neuromuscular diseases (see PMID 23631896, as exemplified by the two cases below.

EXON SKIPPING THERAPY FOR DUCHENNE MUSCULAR DYSTROPHY

Duchenne muscular dystrophy (DMD), a severe and progressive X-linked recessive muscular dystrophy, results from a deficiency of the dystrophin protein. Affected boys need to use wheelchairs by 12 years of age, develop additional cardiomyopathy after age 18 years, and often die before 30 years of age. The disorder is primarily due to intragenic deletions in the 2.4 Mb dystrophin gene.

Surprisingly, deletion of a large central portion of the dystrophin gene, up to 1 Mb, can result in the milder Becker muscular dystrophy (BMD), but deletion of a single nucleotide within an exon in that same 1 Mb central region can result in severe DMD. Two observations explain that apparent anomaly. First, the sequence of the central region of dystrophin is not so important: it acts simply as a flexible linker between the two functionally important parts, the N-terminal and C-terminal domains; large, in-frame internal deletions may reduce dystrophin performance but some functional protein remains, resulting in a mild BMD phenotype. Secondly, internal frameshifting deletions are consistently associated with DMD.

Exon skipping therapy in DMD patients who have a central frameshifting deletion aims to restore the translational reading frame: the net effect should resemble in-frame deletions associated with milder BMD. **Figure 1** shows how the antisense oligonucleotide eteplirsen can induce skipping of exon 51 to restore the translation reading frame in patients with a deletion of exon 50. Skipping of exon 51 can also restore the translational reading frame for several other common internal deletions in the dystrophin gene. Therapeutic skipping of exon 53 has also been carried out using another AO, golodirsen—see Table 9.8. For a review of splicing therapy in neuromuscular disease, see PMID 23631896

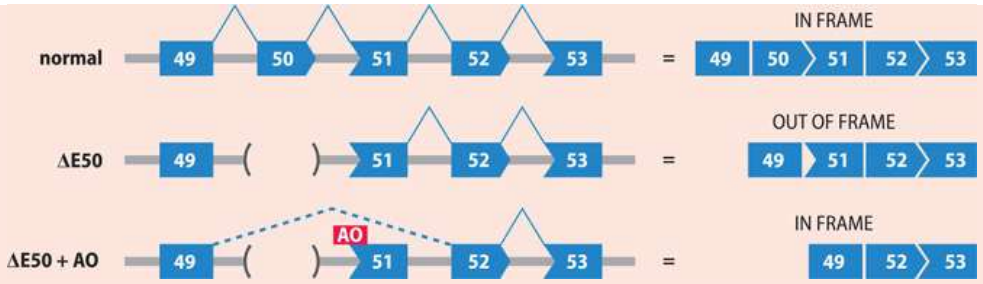


Figure 1 An example of exon skipping therapy for Duchenne muscular dystrophy (DMD). Deletion of the 109-nucleotide exon 50 (ΔE50) in the dystrophin gene results in a shift in the translational reading frame for dystrophin mRNA, resulting in DMD. Local intramuscular injections with eteplirsen, a specific antisense oligonucleotide (AO) that can bind to and blockade splice regulatory sequences at the start of exon 51, causes skipping of exon 51 in ΔE50 patients and splicing of exon 49 to exon 52 without a frameshift (the total number of missing nucleotides = 342). Significant clinical benefit is evident, as measured by improved walking statistics when compared with controls.

TABLE 9.8 THE FIRST WAVE OF MARKETED ANTISENSE OLIGONUCLEOTIDE (AO) AND RNA INTERFERENCE (RNAI) THERAPIES

Type of therapy	Name of therapeutic	Disorder treated	Target gene	PMID
AO gene suppression	inotersen	Hereditary transthyretin amyloidosis	*TTR* (transthyretin)	29972757, 29972750
AO splice modulation	eteplirsen	Duchenne muscular	*DMD* exon 51	29752304
	golodirsen	dystrophy	*DMD* exon 53	32139505
	nusinersen	Spinal muscular atrophy	*SMN2* exon 7	29443664
RNAi (siRNA)	patisiran	Hereditary transthyretin amyloidosis	*TTR* (transthyretin)	29972753, 29972750
	givosiran	Acute hepatic porphyria	*ALAS1* (5-amino-levulinic acid synthase	32521132
	lumasiran	Primary hyperoxaluria type 1	*GO* (glyoxylate oxidase)	33789010

EXON INCLUSION THERAPY FOR SPINAL MUSCULAR ATROPHY

Spinal muscular atrophy (SMA) is a degenerative motor neuron disorder that leads to muscle atrophy and respiratory failure. Individuals with the most severe form rarely survive beyond 2 years of age. The disease is due to defects in the *SMN1* gene which is part of a cluster of duplicated genes that arose by evolutionarily recent segmental duplication. *SMN2*, a paralog

(gene duplicate) of *SMN1*, can produce a protein identical to the SMN1 protein. A single nucleotide change in a splice regulatory sequence, however, causes skipping of exon 7 in 90 % of the *SMN2* transcripts, producing an unstable protein; only 10 % of *SMN2* transcripts make the normal protein (see **Figure 2A**).

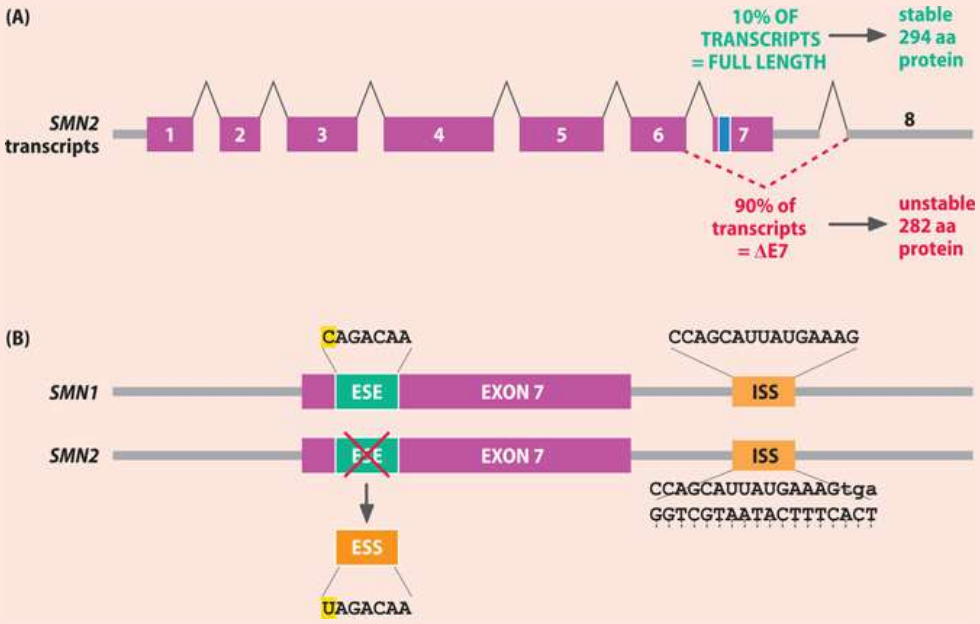


Figure 2 ***SMN2* transcripts, exon 7 splice regulation and role of nusinersen.** 90 % of *SMN2* transcripts lack exon 7 because a 7 bp exonic splice enhancer (ESE) sequence at the beginning of exon 7 in *SMN1* has been mutated; the resulting sequence UAGACAA now works more as a weak exonic splicing suppressor (ESS). Together with an intronic splice silencer (ISS) in intron 7, the effect is to strongly inhibit exon 7 inclusion in the absence of an exon 7 ESE. The antisense oligonucleotide nusinersen works by base pairing with the intron 7 splice silencer sequence, thereby blockading it and promoting exon 7 inclusion in *SMN1* transcripts. Interested readers can find a recent review at PMID 29422644.

The nucleotide sequences of *SMN1* and *SMN2* mRNAs differ at just two nucleotide positions, but one of them, a C/U difference near the beginning of exon 7 is critically important. It falls within a critical exonic splice enhancer sequence, **C**AGACAA in *SMN1*, but the equivalent **U**AGACAA sequence in *SMN2* acts weakly in the opposite direction, as a splice suppressor.

The number of *SMN2* genes can vary as a result of unequal crossover and affected individuals with multiple *SMN2* genes are less severely affected. All individuals with spinal muscular atrophy type 4, the mildest form, have four to six *SMN2* gene copies. *SMN2* may be viewed as a poorly efficient back-up gene, but when there are no functional *SMN1* genes, the more backup *SMN2* genes the better. That prompted the idea of a novel therapy: making the back-up *SMN2* gene more effective by promoting exon 7 inclusion in *SMN2* transcripts using an antisense oligonucleotide, nusinersen (see **Figure 2B**).

RNA interference therapy

Different diseases are potentially amenable to treatment by taking advantage of RNA interference (RNAi), an innate defense mechanism that protects cells against invading viruses and over-active transposable elements. (A small percentage of our resident transposons are actively transposing; if that percentage were allowed to become too great, the genome could be overwhelmed by transposons inserting into essential genes.)

RNAi is triggered in cells by the presence of double-stranded RNA (which is not normally produced in our cells, except by invading viruses, and by the association of sense and antisense transcripts from highly repeated transpo-sons). The double-stranded RNA is detected and cleaved in cells by a ribonuclease called dicer, producing fragments 21 bp long with recessed 5¢ ends, known as **short interfering RNA** (**siRNA**). They are recognized by special protein complexes, RNA-induced silencing complexes (RISC), that initiate a pathway whereby *any* RNA transcripts containing the same nucleotide sequence as the siRNA are destroyed (**Figure 9.20**).

Using RNAi to silence a mutant allele

The pathway shown in Figure 9.20 is concerned with *natural* gene silencing that destroys transcripts from the genes of invading viruses or from transposable elements. It can be artificially exploited to selectively inhibit the expression of a gene of interest within cells. To do that a genetic construct is transfected into cells to produce directly, or indirectly, a gene locus-specific or, preferably an allele-specific, double-stranded siRNA. RISC complexes can then be activated by the allele-specific siRNA to downregulate RNA transcription from, say, a positively harmful gene in the cells of an affected individual. The bulky, highly charged double-stranded siRNA can be transfected into cells with the assistance of attached lipids; alternatively, a gene encoding short hairpin RNA, a siRNA precursor, is transfected into the cells (**Figure 9.21**).

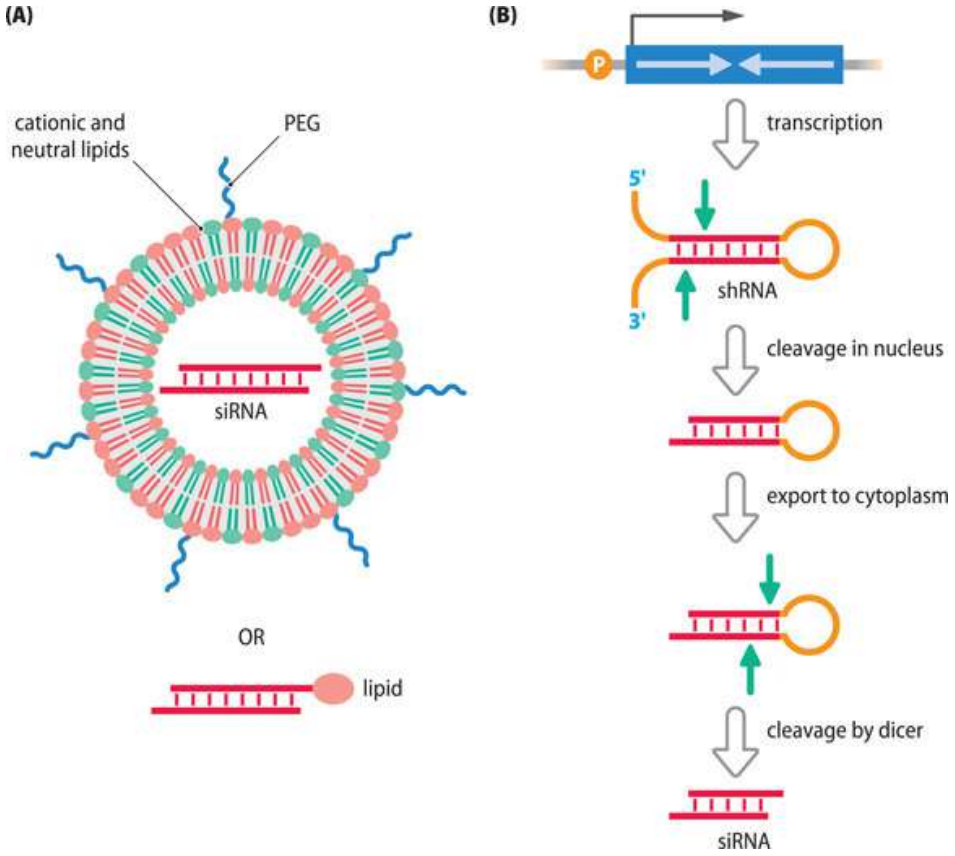


Figure 9.21 Two different types of siRNA delivery to cells. (A) Direct siRNA delivery. The interfering RNA needs to be short because transfecting long double-stranded RNA into mammalian cells results in *indiscriminate* destruction of mRNAs. Two short oligoribonucleotides can be chemically synthesized to form a siRNA duplex that will have two-nucleotide 3¢ overhangs like the natural siRNAs shown in Figure 9.20. The RNA sequence can be chosen to be gene- and allele-specific (a unique sequence from an exon of the target gene that shows differences between mutant and normal alleles). Because siRNA is highly charged and comparatively large, it cannot easily cross plasma membranes; it has to be complexed with lipid-based carriers such as liposomes or conjugated with a lipid such as cholesterol. (B) Delivery of a gene encoding a siRNA precursor. Recombinant viruses are used to ferry an artificial gene construct into cells. The gene has inverted repeats (pale blue arrows) and is transcribed in the nucleus to make a single-stranded RNA with two long complementary sequences, allowing it to fold back to form a mostly double-stranded shorthairpin RNA (shRNA). The shRNA will be processed by the cell's RNAi machinery to yield a specific siRNA duplex in the cytoplasm.

RNAi therapy seeks to silence a specific gene by designing an artificial siRNA to target transcripts of that gene, causing their degradation. It has not been easy: complete gene silencing is difficult to obtain (but significant lowering of the amount of harmful gene product may produce clinical benefit), there is the risk of off-target effects, and efficient delivery to tissues and cells has been problematic. After a series of failures (mostly because of the delivery problem), a corner was turned in 2018: delivery of patisiran, a siRNA specific for the transthyretin gene *TTR*, provided clinical benefit for people with hereditary transthyretin amyloidosis, gaining regulatory approval. The normal transthyretin protein forms a tetramer,

but mutant transthyretin leads to harmful amyloid deposits in different tissues. Affected individuals show slowly progressive peripheral sensorimotor and/or autonomic neuropathy as well as non-neuropathic changes of cardiomyopathy, nephropathy, vitreous opacities, and CNS amyloidosis.

Subsequently, Alnylam, the company responsible for producing transthyretin, developed an effective way of delivering siRNA to liver (by complexing the siRNA with N-acetylgalactosamine, targeting delivery to hepatocytes), and have been researching ways of efficiently targeting other tissues. **Table 9.8** provides a list of RNAi and antisense oligonucleotide drugs that have been approved for treating genetic disorders.

Future therapeutic prospects using CRISPR-Cas gene editing

Genome editing describes any method allowing precise genetic alteration to a pre-determined locus in intact cells. In standard genome editing a first requirement is to have some way of making DNA breaks specifically at the locus of interest. Thereafter, the breaks are exploited in some way in order to obtain a specific desired change to the DNA sequence at a locus of interest. The first such method—*gene targeting*—used endogenous endonucleases that work naturally in homologous recombination; the specificity comes from inserting a transgene carrying a DNA sequence homologous to the gene of interest plus a DNA marker sequence to select for recombination events in which the original sequence was replaced by a desired sequence. The method, detailed in Box 9.2, is rather laborious and time-consuming.

In the newer genome editing methods, genetically engineered constructs are transfected into cells of an affected individual and expressed to make artificial *programmable* endonucleases, ones designed to cut the DNA at pre-determined target sites in the genome. The endonucleases must be transported to the desired target site by being bound to artificially designed RNA or protein **guide sequences** that are specifically designed to bind to the desired target sequences. (Effectively it is the guide sequence that is programmable, being designed to bind a specific sequence, usually 18–20 nucleotides long, in the genome). After a guide sequence has transported its attached nuclease to the correct target sequence, the nuclease cuts the bound DNA strand in the immediate vicinity of the binding site.

CRISPR-Cas: origins

Genome editing using protein guide sequences (such as zinc finger nucleases) is laborious. Happily, however, the CRIPSR-Cas system, which uses RNA guide sequences, is fast, versatile and comparatively simple, and has been transformative. (The acronyms are: CRISPR—**c**lustered, **r**egularly **i**nterspersed **s**hort **p**alindromic **r**epeats; Cas—**C**RISPR-**as**sociated). Like restriction nucleases and RNAi, CRISPR-Cas genome editing was developed from a bacterial self-defense mechanism, in this case a form of adaptive immunity. Here two types of RNA play a critical role:

- guide RNAs, each having a distinctive guide sequence at the 5¢ end (originating from previously captured virus or plasmid sequences) plus a common 3¢ repeat sequence, the R sequence—see **Figure 9.22**.

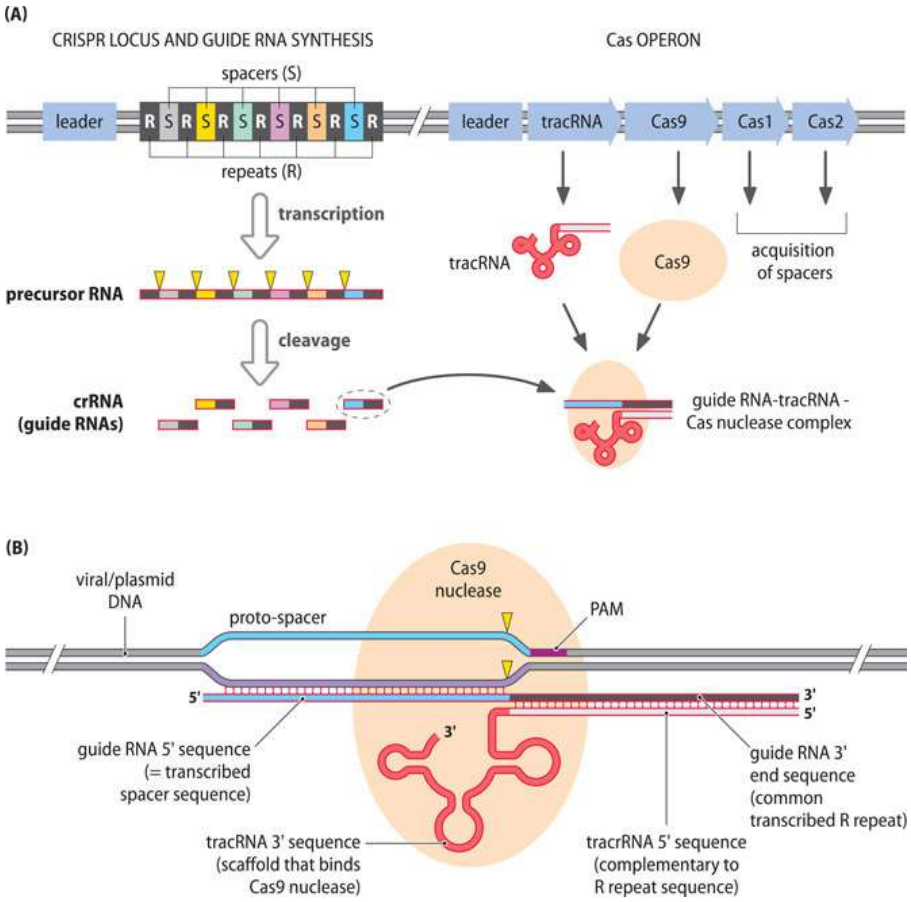


Figure 9.22 How CRISPR-Cas works to defend prokaryotes. (A) A prokaryotic CRISPR locus has multiple copies of an invariant host-cell repeat sequence (R) with interspersed DNA spacers (S), different sequences captured from "proto-spacer" sequences in the genomes of viruses or plasmids that have previously infected the prokaryotic cell. A CRISPR locus produces various short CRISPR *guide RNAs* (gRNAs), each with a distinctive transcribed spacer sequence at its 5¢ end (a *guide sequence*) but a common (invariant) repeat sequence (R) at its 3¢ end. In this example from *S. pyrogenes*, the Cas operon makes a Cas9 (Crispr-associated) nuclease, and a tracRNA (**tr**ans**ac**tivator RNA). The latter works as an intermediate, forming a ternary complex by binding a guide RNA (using its 5¢ end sequence to base pair with the R sequence of a guide RNA), and its 3¢ end sequence to bind a Cas nuclease—see panel B for an expanded view. (B) Defence against recurring virus or plasmid invasion. The blue proto-spacer sequence on the invading DNA is identical to a spacer previously captured and stored in a CRISPR locus (the spacer in the dashed ring on the right of the CRISPR locus in panel A). The 5¢ end of an appropriate guide RNA (as part of a guide RNA-tracRNA-Cas nuclease complex) can therefore bind to a complementary sequence on the invading DNA; the binding occurs just upstream of a short protospacer-associated motif (PAM) in the virus/plasmid DNA (which in the

case of the Cas9 nuclease is 5¢ NGG 3¢, where N = A, C, G or T). Once the Cas9 nuclease has been brought close to the target viral/plasmid DNA it cuts the DNA on both strands (vertical yellow darts).

- a tracRNA (***tr****ans*-**ac**tivating RNA) having at the 5¢ end a sequence complementary to the R sequence, and at the 3¢ end a sequence that can act as a scaffold to specifically bind a CRISPR-associated (Cas) nuclease—see Figure 9.22.

Exploiting CRISPR-Cas for therapeutic genome editing

CRISPR-Cas genome editing began by exploiting the natural CRISPR-Cas system to make double-stranded DNA breaks in the genomes of complex cells. Synthetic guide RNAs would be designed to recognize target sequences 18–21 nucleotides in length. After a double-strand break would be made at the desired locus, cellular DNA repair mechanisms would be activated that could produce desired sequence changes at the gene of interest, or be manipulated to do so. Recall from Section 4.2 that two DNA repair pathways are dedicated to repairing double-strand DNA breaks in our cells, and they can be exploited in genome editing as detailed below.

- Nonhomologous end joining (NHEJ) is commonly used by cells, operates throughout the cell cycle and prioritizes speed over accuracy. It rapidly joins the ends of a broken DNA but very small mistakes are made during repair (a nucleotide is often deleted or inserted, for example). When carrying out *ex vivo* gene therapy, cells subjected to genome editing can be sampled to see if, in some of them, NHEJ produces the desired sequence change. This type of DNA repair can be used to inactivate a functional sequence for a therapeutic purpose.
- Homology-dependent DNA repair. The homologous recombination DNA repair pathway is available to replicating cells after S phase. When used naturally by cells it can make accurate repairs to a double-strand DNA break using the unbroken sister chromatid as a DNA template. A type of homologous recombination can be used to repair a pathogenic mutation by engineering a double-strand break in a mutant allele and simultaneously providing a transgene containing a sequence from a normal allele and flanking sequences with 100 % sequence homology. This procedure, known as *homology-directed repair (HDR)*, can repair the mutant allele, converting the sequence of a disease allele to that of a normal allele (**Figure 9.23**).

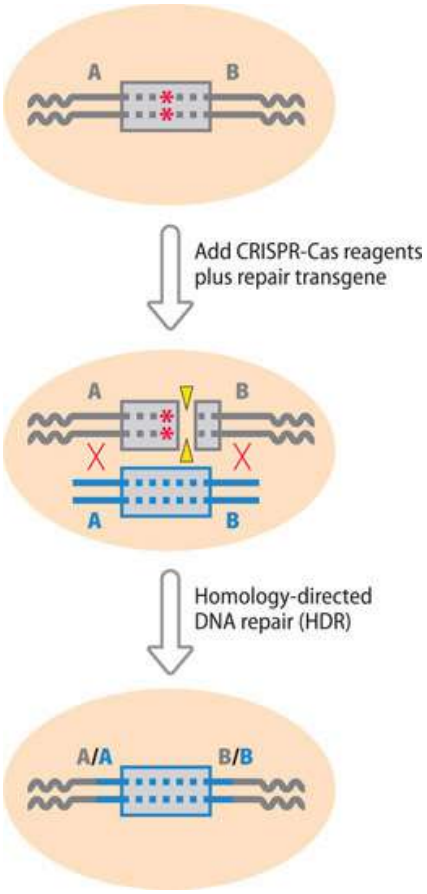


Figure 9.23 Homology-directed DNA repair A chromosomal gene with a pathogenic mutation (shown here as a red asterisk within an exon) can be repaired by using CRISP-Cas to make a double-stranded break in the immediate vicinity of the mutation. A provided transgene (blue), with the normal DNA sequence for the exon and flanking intronic sequence A and B, can act as a template. Recombination (marked by red X) between the flanking sequences of chromosomal exon and the transgene can replace the mutant sequence by a normal sequence. Note that for point mutations, a short single-stranded oligonucleotide is often preferentially used as a template DNA.

A single hybrid RNA (sometimes called sgRNA) is commonly used in modern CRISPR-Cas genome editing, with a 5¢ guide sequence from a guide RNA joined by a linker sequence to the 3¢ scaffold sequence of tracRNA. The target sequence must have a suitable protospacer motif (PAM) immediately downstream of it. In modern CRISPR-Cas genome editing, modified Cas nucleases, called *DNA nickases*, are often used that are designed to cut a *single* DNA strand, and two closely neighboring target sequences are often designed to be bound, as shown in **Figure 9.24**. That has two important consequences. First, it minimizes inappropriate binding of the guide sequence to *off-target* sequences (any sequence other than the desired target sequence). Secondly, the accuracy in making a desired change may be improved: errors are common in artificial homology-directed DNA repair when both DNA strands are broken.

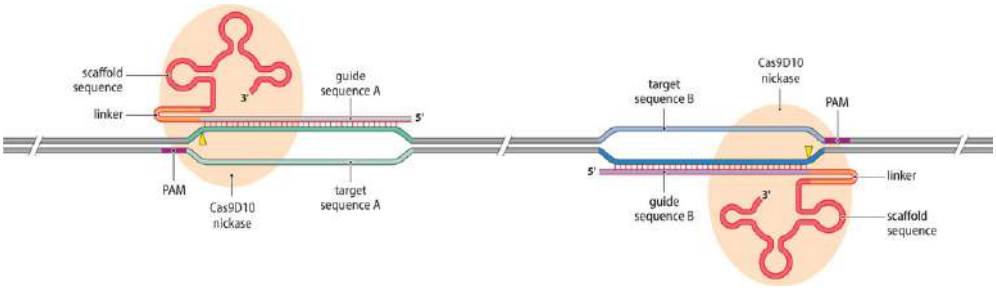


Figure 9.24 Making single-strand DNA breaks at neighboring target sequences using DNA nickases in CRISPR-Cas genome editing. The object is to reduce off-target effects by delivering a modified Cas nuclease to *two* closely located target sequences, A and B, at a locus of interest. The modified Cas nuclease is a DNA nickase, able to cut just one of the two DNA strands (at positions shown by the vertical yellow darts). The DNA nickases are delivered to their target sequence by a single type of RNA (containing the guide sequence from a guide RNA attached via a linker sequence to a tracRNA scaffold sequence that binds the DNA nickase).

Subsequent variants of the standard CRISPR-Cas design include methods where Cas nucleases are fused to proteins with some type of enzyme activity. Two such methods are listed below.

- *Base editing.* A catalytically impaired Cas nuclease is fused to an enzyme that converts one base to another, without the need for cutting the DNA or for a DNA template. Initial work involves cytosine base-editors, such as the cytosine deaminase APOBEC (converts C to U; the U is subsequently converted to T after DNA replication or DNA repair) and adenine base-editors. This method may be superseded by the one directly below.
- *Prime editing. A Cas* nuclease is fused to a reverse transcriptase, and a special prime-editing guide RNA (pegRNA) is used. In addition to the usual guide sequence and Cas-binding scaffold sequence, the pegRNA has an additional "replace" sequence; one that acts as a template for inserting a short desired sequence. Because it offers precise genome editing, and can carry out any type of short sequence correction, this may become a very popular method.

At the time of writing (January 2022) it is still early days for therapeutic applications of genome editing, but the technology is advancing rapidly. Some clinical trials have used older-style technologies featuring protein guide sequences, but now the focus is very much on CRISR-Cas genome editing.

As in the case of standard supplementation gene therapy, *ex vivo* gene editing therapies are primarily being used because of the huge advantage of being able to study autologous cells in culture and then selecting those that show the desired genetic modification before injecting the modified cells into a patient. Gene editing therapies may offer some advantages over gene supplementation therapies because of the size limitation of insert DNA that can be accommodated in gene therapy vectors. For certain genes, a cDNA may simply be too large to

be accommodated in vectors, notably AAV vectors. But it too has some downsides—notably the potential problem for immune reactions when expressing the Cas nuclease.

Genome editing strategies that rely on NHEJ (nonhomologous end joining) DNA repair are inherently more efficient than HDR (homology-directed DNA repair), and are being pursued where the object is to inactivate a gene or a regulatory sequence (see **Table 9.9** for examples).

TABLE 9.9 SOME EARLY EXAMPLES OF USING GENOME ENGINEERING IN NON-CANCER CLINICALTRIALS

Type of approach	Disorder treated and basis of method	Technology and reference
Disable a cell receptor to prevent virus infection	HIV-AIDS. Inactivate the gene making the CCR5 receptor on helperTcells (required for HIV infection) by the HIV virus. *Ex vivo* gene therapy using autologous$CD34^{+}$T cells.	ZFN, TALEN, CRISPR
Alter regulatory signals so as to reactivate a silenced gene to make a protein to supplement genetic deficiency of a closely related protein	Beta thalassemia/sickle cell disease. The idea is to restore gamma globin production (normal in fetal stages only) to make up for lack of normal beta globin in affected individuals. May involve targeting the cis-acting BC11A repressor of the gamma-globin gene or its target sequence.	ZFN, CRISPR (see Figure 9.3 ofPMID 32775490)

ZFN (zinc finger nuclease) and TALEN (TALE nuclease) are older, cumbersome genome editing technologies that use protein guide sequences.

Therapeutic applications of stem cells and cell reprogramming

As detailed earlier, many of the successes in gene supplementation therapy, notably *ex vivo* gene therapies, have been dependent on cell transplantation and therefore also constitute a type of cell therapy. But stem cells also offer the prospect of **regenerative medicine**, a type of *cell supplementation therapy* in which stem cell cultures are manipulated so as to simply provide replacement cells for cells lost through disease (or injury). Take cultured human pluripotent stem cells, which can proliferate indefinitely and differentiate into all types of cells in the body. If efficiently directed down the correct differentiation pathway, they could in principle provide replacement cells to supplement a deficiency of some functioning cells in a patient. Complex disorders arising from loss of a particular cell type, are possible targets for this type of therapy, including Type I diabetes and Parkinson disease (loss of pancreatic beta cells and dopaminergic neurons, respectively), as are some injuries (for example, to the spinal cord).

Sources of cells for cell therapy

Most cell therapies have used cultured human pluripotent stem cells. Of these, human embryonic stem cells (ESCs) were the first to be obtained (by culturing surplus human embryos from *in vitro* fertilization centers) and have been with us for more than 20 years. More recently, human somatic cells have been induced to dedifferentiate to produce **induced pluripotent stem cells** (**iPSCs** – see Box 9.1), which are more acceptable ethically than ESCs and are now being used in preference to ESCs for therapeutic applications. iPSC technology also has the advantage of permitting, in principle, *ex vivo* genetic modification to be extended to a wide range of genetic disorders, not just disorders of blood or of other cells originating from hematopoietic stem cells. (In this case, iPSCs derived from accessible blood or skin cells in an affected individual would first be genetically modified in culture and differentiated to give a desired cell type; then, the desired, genetically modified, differentiated cells would be returned to an appropriate location in the patient.)

Cells obtained after directed differentiation of ESCs or iPSCs have been used in clinical trials to treat various disorders, including various eye disorders and some complex diseases. But this is still a young field and although there is considerable promise, various obstacles need to be surmounted, as explained below. Another possible source of cells involves *transdifferentiation*, switching from one differentiated cell type to another, such as from fibroblast to neuron. The method is technically difficult, but one interesting possibility of a therapeutic application is to convert astrocytes into neurons *in vivo* to treat Parkinson disease. Astrocytes, a subtype of glial cells, are the most abundant cells in the central nervous system and can be converted into induced dopamine-releasing neurons by blocking expression of an astrocyte protein called PTB. Interested readers can find details of preliminary work using a mouse model of Parkinson disease at PMID 32581373.

Obstacles to overcome in cell therapy

Three major types of obstacle need to be surmounted in cell therapies, as shown in **Figure 9.25A**. Of these immunogenicity is a major problem. Any transplant of cells runs some risk of immune rejection: cell surface antigens on transplanted cells, most notably variant HLA proteins, may be perceived as foreign. (Note, however that the degree of immune response is less in some sites, including the central nervous system and most of the the eye, which have *immune privilege*: foreign antigens may be tolerated without inducing an inflammatory immune response.) Cells originating from differentiation of ESCs, which derive from donors of surplus embryos, and so are allogeneic, need to be transplanted into a patient where there is a high degree of matching for HLA antigens. Although transplant of *autologous* iPSC-derived cells back into the patient of origin is not normally expected to provoke immune reactions, the problem here is the huge expense involved in making autologous iPSCs from individual patients.

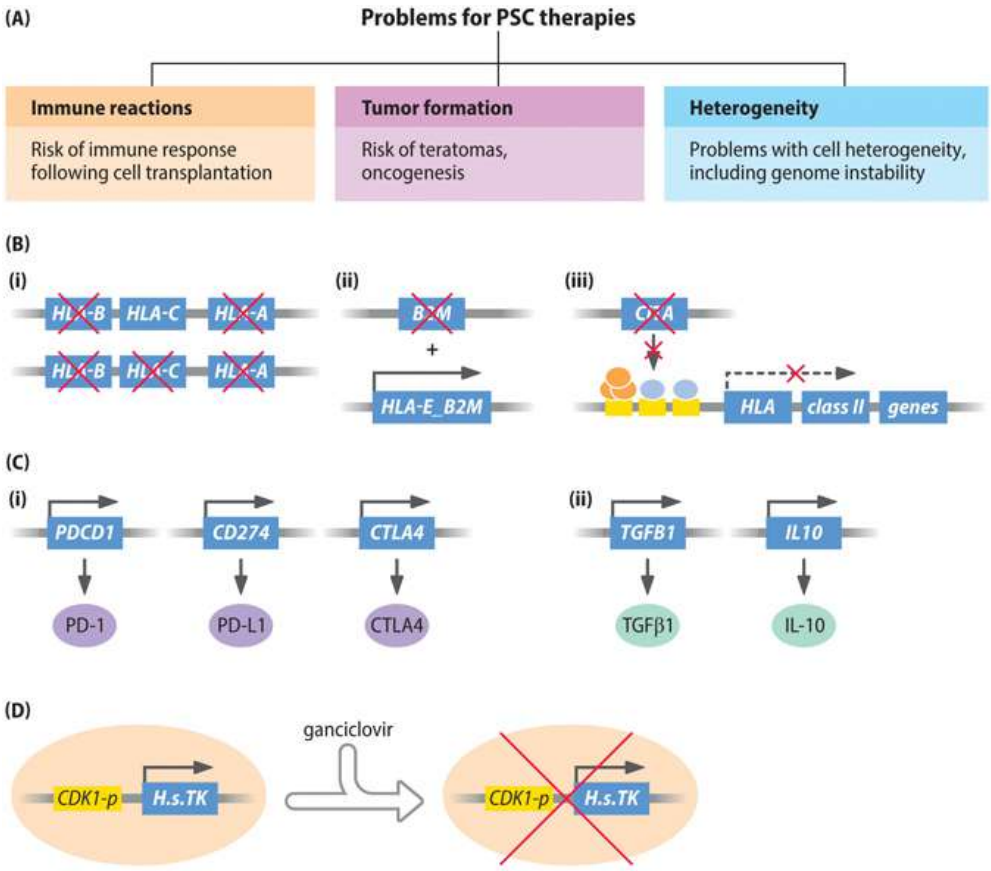


Figure 9.25 Overcoming problems for therapeutic application of pluripotent stem cells (PSCs). (A) The three main difficulties to overcome. (B) Suppressing HLA protein production by genome editing. Immunogenicity can be reduced by deleting five of the six polymorphic class I HLA genes leaving a single, mildly polymorphic HLA-C allele, as shown in B(i), or by deleting the *B2M* gene and supplying a fusion gene combining the non-polymorphic *HLA-E* gene with a *B2M* gene copy, as shown in B(ii). Class II HLA protein production can be switched off by deleting the *CIITA* gene that makes an activator protein required for class II HLA expression—see B(iii). (C) *Immune cloaking.* This can be done by stimulating (i) genes making proteins involved in checkpoint blockade; and (ii) genes making localized immune suppressants. (D) Protecting against teratoma formation. After introducing into cells a suicide gene, the *Herpes simplex* thymidine kinase (*TK*) gene, coupled to a *CDK1* gene promoter (always expressed in dividing cells), treatment with ganciclovir selectively kills dividing cells that express the *TK* gene.

To address the immunogenicity problem, banks of different iPSC lines have been set up in some countries from rare donors homozygous for frequently occurring HLA haplotypes. According to the degree of HLA matching, an optimal iPSC line could be selected for use with a patient to reduce the chance of immune reaction. An alternative approach is to apply genome editing to iPSCs to make them "immunocompatible" in some way.

Another way of making cells immunocompatible is to suppress expression of polymorphic HLA proteins. The classical class I HLA genes, notably *HLA-A* and *HLA-B*, and to a lesser extent *HLA-C*, make highly polymorphic heavy chains that each combine with an invariant

light chain, beta-2-microglobulin, produced by the *B2M* gene. If genome editing simply deletes the *B2M* gene, the genetically modified cells have no class I HLA antigen; although no longer detected by cytotoxic T cells, these "unnatural" cells are liable to be killed by natural killer cells. To counter natural killer cells, genome editing seeks to leave some residual modestly or poorly polymorphic HLA protein (Figure 9.25Bi,ii). Class II HLA expression can be suppressed by deleting the *CIITA* (Class II major histocompatibility complex transactivator) gene (Figure 9.25Biii).

A different way of reducing the immunogenicity of therapeutic cells is *immune cloaking*, which involves stimulating the expression of genes that cancer cells use to escape immune detection (Figure 9.25C). The ultimate aim is to generate "universal" hypoimmunogenic pluripotent stem cell lines to be used as "off-the-shelf" reagents; differentiated cells derived from them could be transplanted into any patient with minimal risk of immune reaction.

Another important obstacle in cell therapy is the possibility of tumorigenesis. Practical difficulties in accurately and efficiently directing iPSCs (or ESCs) to undergo differentiation steps towards a desired differentiated cell type could lead to incomplete differentiation. Residual pluripotent cells might then be transmitted to the patient that could form *teratomas*, tumors composed of heterogeneous cell types derived from the different embryonic germ layers. One way to safeguard against that is to genetically modify cells derived by iPSC differentiation so that they contain a *suicide gene* that is expressed only in dividing cells (Figure 9.25D).

A special case: preventing transmission of severe mitochondrial DNA disorders by mitochondrial replacement

Mutations in mitochondrial DNA (mtDNA) are a significant cause of human disease: pathogenic mutations are found in at least 1 in 200 of the population, and cause severe multisystem disease in approximately 1 in 10 000 of the population. Pathogenic mtDNA can be maternally inherited, but there are no effective treatments for mitochondrial DNA disorders.

In the clinical management of mtDNA disorders, the emphasis has therefore been on prevention. Preimplantation and prenatal diagnosis (as described in Chapter 11) are well established in clinical genetic practice as a way of selecting unaffected embryos. However, the results can be difficult to interpret for patients with *heteroplasmic* mtDNA mutations (with variable numbers of mutant and normal mtDNAs in each cell). In addition, an increasingly large group of diseases are recognized to be caused by *homoplasmic* mtDNA mutations (all, or almost all, of the mtDNA molecules are mutant). Here, prevention by selecting an unaffected embryo is not an option—all the offspring would inherit the pathogenic mutation in the maternal egg, and this type of genetic defect can be associated with a very high disease recurrence risk.

Women who are carriers of serious mtDNA disorders caused by homoplasmic mutations or where the proportion of mutant mtDNA is close to 100 % therefore face the bleak prospect of

having severely affected children in *each* pregnancy. The usual option of prenatal diagnosis to select healthy embryos cannot be achieved if every embryo will contain mutant mtDNA.

The transmission of homoplasmic mutations can, however, be avoided if the defective maternal mtDNA is replaced by mtDNA from an asymptomatic donor. That can be done by *in vitro* fertilization using either of two slightly different approaches, replacing mtDNA at the zygote level or the oocyte level (**Figure 9.26**). The resulting human embryos appear to be viable *in vitro*, and the degree of mutant DNA carryover is low or undetectable. The clinical application of this mitochondrial replacement method became legally permissible in the United Kingdom in 2015, and is now a nationally commissioned NHS services in England and Wales. It constitutes an exceptional example of germline gene transfer in humans, and we return to consider the associated ethical considerations in Chapter 11.

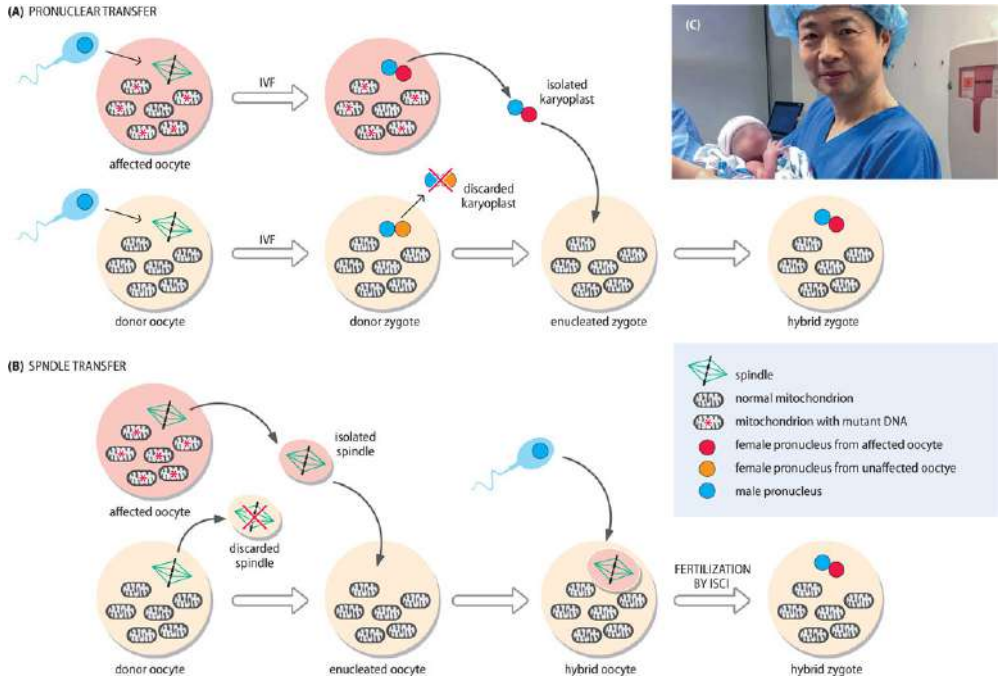


Figure 9.26 Mitochondrial replacement therapy to prevent transmission of severe mtDNA disease. A donor provides an enucleated oocyte with healthy mitochondria and normal mtDNA; the prospective parents provide the nuclear genome, either after or before *in vitro* fertilization (IVF). (A) Pronuclear transfer technique. An affected oocyte from the prospective mother (with many, sometimes all, mtDNAs having the pathogenic mutation) is fertilized by her partner's sperm. The resulting normal *karyoplast* (combined male and female pronuclei) is isolated, then transferred into an enucleated donor zygote with normal mitochondria. The resulting zygote has "foreign" but normal mtDNA. (B) Metaphase II spindle transfer technique. The metaphase II spindle is transferred from an oocyte with mutant mtDNA into a mitochondrial donor oocyte. The resulting hybrid oocyte has a nuclear genome from the prospective mother, but mtDNA from the donor. Fertilization by intracytoplasmic sperm injection (ICSI) produces a hybrid zygote. (C) In an attempt in 2016 to prevent transmission of Leigh syndrome, a severe neurological disorder, a hybrid human zygote produced by mitochondrial donation gave rise to a three-parent baby. The image shows the healthy baby boy. When tested some months later just 1 % of his mtDNA carried the harmful mutation. Holding him is Dr. John Zhang from

the New Hope Fertility Center in New York. (A and B adapted from Craven L et al. [2011] *Hum Mol Genet* 20:R168–174; PMID 21852248, with permission from Oxford University Press; C reproduced courtesy of New Hope Fertility Center, New York.)

SUMMARY

- Treatment for inborn errors of metabolism sometimes involves supplementing a genetic deficiency, but often the treatment is directed at reducing the harmful effects of abnormally elevated metabolites.
- Drug development typically involves screening hydrocarbon-based small molecules for compounds that will bind to medically important protein targets. By binding to a protein, the drug affects its function in some way.
- Genetic variation means that different individuals can respond very differently to drugs; adverse reactions to drugs are very common and cause very many fatalities.
- The pharmacokinetics of a drug describes how it is absorbed, activated (in the case of a prodrug), metabolized, and excreted; pharmacodynamics describes the effect it has on the body.
- Phase I drug metabolism reactions are typically oxida tive reactions carried out by monooxygenases; phase II reactions are conjugative reactions in which a transfer-ase enzyme adds a chemical group. The overall effect is to convert lipophilic hydrocarbon drugs into more polar forms that can be excreted more easily.
- In addition to dealing with artificial drugs, drug-metabolizing enzymes handle unusual exogenous chemicals (xenobiotics) in our diet and environment. They are often highly polymorphic because xenobiotics originating from other organisms are under genetic control and potentially harmful to us.
- The therapeutic window is the range of drug concen trations in which pharmaceutical benefit is achieved without safety risks.
- Poor drug metabolizers are at risk of a drug overdose (the drug does not get cleared quickly; repeated drug doses drive up the concentration). Others are ultrafast metabolizers and may get little therapeutic benefit (the drug is cleared too rapidly).
- Six cytochrome P450 enzymes carry out 90 % of phase I drug metabolism. Each handles the metabolism of multiple drugs; conversely, some individual drugs may be metabolized by two or more cytochrome P450 enzymes.
- When a drug is metabolized principally by one enzyme, genetic variation in that enzyme can be mostly responsible for large differences between individuals in

the ability to metabolize that drug. For some other drugs, such as warfarin, several different genetic factors determine how the drug is metabolized.

- Therapeutic "recombinant proteins" are made by expressing cloned human genes in cells to make a human protein that can be purified and used to treat a genetic deficiency of that protein.
- Therapeutic antibodies are usually designed to bind to harmful gene products to block their effects. Rodent monoclonal antibodies have limited lifetimes after injection into patients; genetic engineering allows the replacement of rodent sequences by human sequences to make more effective antibodies.
- Genetically engineered antibodies with a single vari able polypeptide chain can work as intracellular anti bodies by binding harmful proteins within cells.
- Gene therapy means inserting nucleic acids or oligo-nucleotides into the cells of a patient to counteract or alleviate disease.
- In gene supplementation therapy, diseased cells that are genetically deficient for some product are supplemented by transfecting a cloned gene to make the missing product inside the cells.
- Some therapies target RNA. In gene silencing, the expression of a positively harmful gene (such as a gene with a gain-of-function mutation or one expressed by a pathogen) is selectively repressed, usually by inhibiting the RNA. RNAs can sometimes also be induced to undergo alternative splicing to counteract disease.
- Stem cells are cells that can both renew themselves and give rise to more differentiated (more specialized) cells. Pluripotent embryonic stem cells are artificially cultured cells derived from the very early embryo that can be induced to give rise to virtually any differentiated cell. Somatic stem cells help to replace a limited set of short-lived cells.
- In cell reprogramming, the epigenetic settings of cells are artificially altered to induce changes in gene expression so that the cells acquire the characteristics of a different cell type. Differentiated cells can be induced to dedifferentiate to become unspecialized pluripotent stem cells or to form a different type of somatic cell (transdifferentiation).
- Virus vectors are more efficient but less safe than non viral vectors in transporting therapeutic genetic constructs into cells.
- Integrating virus vectors such as lentivirus vectors can allow a genetic construct to be inserted into the chromosomes of a cell. That is highly desirable when targeting short-lived cells that are replenished by stem cells; if a therapeutic transgene integrates into the stem cell, it will be transmitted by cell division.
- *Ex vivo* gene therapy involves removing cells from a patient, genetically modifying them in culture and returning the genetically modified autologous

cells to the patient. It has been used to treat disorders by genetic modification of impure populations of hematopoietic stem cells that give rise to blood cells or some types of tissue immune system cell.

- In *in vivo* gene therapy, the cells of a patient are geneti cally modified *in situ.* Non-integrating virus vectors such as AAV virus are commonly used to transfect differentiated cells.
- Animal disease models are usually created by geneti cally modifying the germ line to mimic a human phenotype. They are important in permitting a detailed understanding of molecular pathology, and to provide a front-line system for testing new therapies.
- Therapeutic antisense oligonucleotides are designed to base pair with RNA transcripts so as to inhibit the expression of harmful gene products of the target gene or to modulate RNA splicing.
- RNA interference (RNAi) is a natural gene-silencing mechanism that evolved as a cellular defense against virus attack or excessive transposon activity. Therapeutic short interfering RNAs can be designed to inhibit expression of a harmful gene or allele after delivery into the cells of a patient.
- Genome editing involves making a precise genetic modification at a pre-determined location in the genome of intact cells.
- Standard CRISPR-Cas genome editing uses artificial hybrid RNAs having a variable guide sequence (programmed to base pair to a desired unique genomic site) and a scaffold sequence (binds to Cas endonuclease). The bound Cas nuclease is transported to the desired site, where it cuts the DNA to permit DNA repair and the desired genetic modification.
- Mitochondrial replacement is an *in vitro* fertilization method that avoids transmission of a severe mitochondrial DNA disorder to a child by moving the nuclear DNA from a maternal oocyte (either prior to, or after, fertilization) to an enucleated donor oocyte (before or after fertilization) which has normal healthy mitochondria, and mtDNA.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

General overviews

Dietz H (2010) New therapeutic approaches to mendelian disorders. *N Engl J Med* 363:852–863; PMID 20818846.

Pharmacogenetics and pharmacogenomics

Cytochrome P450 Drug Interaction Table. *Indiana University School of Medicine, Division of Clinical Pharmacology*. Available at https://drug-interactions.medicine.iu.edu/Home.aspx

Hockings JK (2020) Pharmacogenomics: an evolving clinical tool for precision medicine. *Cleveland Clinic J Med* 87:91–99; PMID 32015062.

Meyer UA, Zanger UM & Schwab M (2013) Omics and drug response. *Annu Rev Pharmacol Toxicol* 53:475–502; PMID 23140244.

Pharmacogenomics Knowledge Base (PharmGKB). http://www.pharmgkb.org [A knowledge resource offering clinical information including dosing guidelines and drug labels, potentially clinically actionable gene-drug associations and genotypephenotype relationships.]

Roden DM (2019). Genomic medicine 2. Pharmacogenomics. *Lancet* 394:521–532; PMID 31395440.

Wang L, McLeod HL & Weinshilboum RM (2011) Genomics and drug response. *N Engl J Med* 364:1144–1153; PMID 21428770.

Small molecule drug therapy for genetic disorders

Curatolo P & Moavero R (2012) mTOR inhibitors in tuberous sclerosis complex. *Curr Neuropharmacol* 10:404–415; PMID 23730262.

De Boeck K & Davies JC (2017) Where are we with transformational therapies for patients with cystic fibrosis? *Curr Opin Pharmacol* 34: 70–75; PMID 28992608.

Franz DN (2013) Efficacy and safety of everolimus for subependymal giant cell astrocytomas associated with tuberous sclerosis complex (EXIST-1): a multicenter, randomized, placebo-controlled phase 3 trial. *Lancet* 381:125–132; PMID 23158522.

Heilbron K (2021) Advancing drug discovery through the power of the human genome. *J Pathol* 254:418–429; PMID 33748968.

Santos R (2017) A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* 16:19–34; PMID 27910877.

The Therapeutic Targets Database at http://db.idrblab.net/ttd/

Therapeutic antibodies and proteins

Cardinale A & Biocca S (2008) The potential of intracellular antibodies for therapeutic targeting of protein-misfolding diseases. *Trends Mol Med* 14:373–380; PMID 18693139.
Dimitrov DS (2012) Therapeutic proteins. *Methods Mol Biol* 899:1–26; PMID 22735943.
Hunter P (2019) The prospects for recombinant proteins from transgenic animals. *EMBO Reports* 20:e48757; PMID 30397525.
Kim SJ (2005) Antibody engineering for the development of therapeutic antibodies. *Mol Cells* 20:17–29; PMID 16258237.
Lu R-M (2020) Development of therapeutic antibodies for the treatment of diseases. *J Biomed Sci* 27:1 PMID 31894001.

Animal disease models for testing therapies

Shen H (2013) Precision gene editing paves way for transgenic monkeys. *Nature* 503:14–15; PMID 24201259.
Strachan T & Read AP (2019) *Human Molecular Genetics*, 5th ed. Garland Science. [Chapter 21 gives a detailed account of the technologies involved in making animal models and the extent to which the phenotypes faithfully replicate that of human disorders they were intended to mimic.]

Gene therapy: general

Anguela XM & High KA (2019) Entering the modern era of gene therapy. *Annu Rev Med* 70:273–288; PMID 30477394.
Dunbar CE (2018) Gene therapy comes of age. *Science* 359(6372):eaan4672; PMID 29326244.
Gene Therapy *Net.com*. Available at http://www.genetherapynet.com/ [Covers basic science, clinical trials and includes databases, publication and so on.]
Ginn SL (2018) Gene therapy clinical trials worldwide to 2017—an update. *J Gene Med* e3015; PMID 29575374.
High KA & Roncarolo MG (2019) Gene therapy. *N Eng J Med* 381:455–464; PMID 31365802.
Ma C-C (2020) The approved gene therapy drugs worldwide from 1998 to 2019. *Biotechnol Adv* 40:107502; PMID: 31887345

Clinical trials databases

ClinicalTrials.gov. www.clinicaltrials.gov [*A comprehensive US government site*.]
Gene Therapy Clinical Trials Worldwide. https://a873679.fmphost.com/fmi/webd/GTCT [Provided by the Journal of Gene Medicine and published by Wiley.]

RNA and oligonucleotide therapeutics

Bennett CF (2019) Therapeutic antisense oligonucleotides are coming of age. *Annu Rev Med* 70:307–321; PMID 30691367.

Crooke ST (2018) RNA-targeted therapeutics. *Cell Metab*. 27:714–733; PMID 29617640.

Dowdy SF (2017) Overcoming cellular barriers for RNA therapeutics. *Nature Biotechnol* 35:222–229; PMID 28244992.

Hu B (2020) Therapeutic siRNA: state of the art. *Signal Transduct Target Ther* 5:101; PMID 32561705

Lieberman J (2018) Tapping the RNA world for therapeutics. *Nature Struct Mol Biol* 25:357–364; PMID 29662218.

Kuijper EC (2020) Opportunities and challenges for antisense oligonucleotide therapies. *J Inher Metab Dis* (in press); PMID 32391605.

Yu A-M (2020) RNA drugs and RNA targets for small molecules: principles, progress, and challenges. *Pharmacol Rev* 72:862-898; PMID 32929000.

CRISPR-Cas and therapeutic genome editing

Doudna JA (2020) The promise and challenge of therapeutic genome editing. *Nature* 578:229–236; PMID 32051598.

Doudna JA & Charpentier E (2014) The new frontier of genome engineering with CRISPR-Cas9. *Science* 346(6213): 1258096; PMID 25430774.

van Haasteren J (2020) The delivery challenge—fulfilling the promise of therapeutic genome editing. *Nature Biotechnol* 38:845–855; PMID 32601435.

Stem cells and cell therapy

Cossu G (2018) *Lancet* Commission: stem cells and regenerative medicine. *Lancet* 391:883–910: PMID 28987452.

Kimbrel EA & Lanza R (2020) Next-generation stem cells—ushering in a new era of cell-based therapies. *Nature Rev Drug Discov* 19:463–479; PMID 32612263.

Lanza R (2019) Engineering universal cells that evade immune detection. *Nature Rev Immunol* 19:723–733; PMID 31417198.

Yamanaka S (2020) Pluripotent stem cell-based cell therapy—promise and challenges. *Cell Stem Cell* 27:523–531; PMID 33007237.

Treatment for mitochondrial DNA disorders

Chinnery PF (2020) Mitochondrial replacement in the clinic. *N Eng J Med* 382:1855–1857; PMID 32374967.

Russell OM (2020) Mitochondrial diseases: hope for the future. *Cell* 181:168–188; PMID 32220313.

10
Cancer genetics and genomics

DOI: 10.1201/9781003044406-10

CONTENTS

Why should cancer merit a separate chapter in this book and not, say, neurology or cardiology? Well, the molecular pathogenesis in cancers is, for the most part, quite different from that of other genetic disorders: somatic mutations and epi-genetic dysregulation are extremely common, and natural selection operates at the level of the cell as well as at the level of the organism. In addition, a number of specialized genetic mechanisms — kataegis, chromothripsis, chromoplexy and so on—are observed only in cancer cells.

Tumorigenesis involves an extraordinary and bewildering degree of changes to both the genome and the epigenome. Not only is there genetic heterogeneity

between tumors, but also within individual tumors. Despite the heterogeneity of the very many different diseases that we call cancer, the phenotype—uncontrolled cell proliferation—is comparatively simple and more amenable to genetic analysis than some other common classes of disease, such as psychiatric disorders.

In Section 10.1 we give an overview of the primary distinguishing biological capabilities of cancer cells, outline the broad multi-stage evolution of cancers and describe how intratumor heterogeneity evolves. Section 10.2 is mostly devoted to considering the principles underlying two fundamental classes of genes in cancer development: oncogenes and tumor suppressor genes. As cancers evolve, genomic instability and epigenetic dysregulation become increasingly prominent; we consider selected aspects in Section 10.3.

Genome-wide molecular profiling studies—notably genome-wide sequencing—are transforming our understanding of cancer, and in Section 10.4 we take a look at new insights emerging from the burgeoning cancer genome studies. Finally, in Section 10.5 we consider the challenges and prospects in deriving clinical benefit from all the extraordinary information coming out of cancer genetics and cancer genome studies.

10.1 FUNDAMENTAL CHARACTERISTICS AND EVOLUTION OF CANCER

In order to appreciate why cancers are so different from other genetic disorders, it is important to understand how cancers evolve and the role of natural selection in this process. First, however, we provide a summary of the fundamental characteristics of cancers.

The defining features of unregulated cell growth and cancer

The term **cancer** is applied to a heterogeneous group of disorders whose common features are uncontrolled cell growth and cell spreading; abnormal cells are formed that can invade adjacent tissues and spread to other parts of the body through the blood and lymph systems (but see below for a second usage). *Carcinogenesis,* the general process of cancer formation, may result from aberrant functioning of a wide range of genetic control mechanisms, as detailed below.

Aberrant regulation of cell growth results in an abnormal increase in cell numbers; growths can result that appear normal or abnormal. A growth containing

excessive numbers of cells that appear to be virtually the same as those in the normal tissue is said to be *hyperplastic;* a growth that has cytologically abnormal cells is said to be *dysplastic*.

Sometimes a growth formed by excessive cell proliferation is localized. That is, it shows no signs of invading neighboring tissue, and is described as a **benign tumor**. Benign tumors are self-limiting: they grow slowly and can often be surgically removed with low risk of recurrence. They often do not present much danger. Sometimes, however, they grow quite large over time, and simply by expanding, they can press on neighboring structures in a way that can cause disease. For example, in tuberous sclerosis complex (caused by mutations in *TSC1* or *TSC2,* genes that work in the mTOR growth signaling pathway), benign tumors usually form in multiple different organs. By growing to a large size, they can sometimes disrupt organ function.

In the more than 100 different diseases that we call cancers, the abnormal cells resulting from uncontrolled cell growth have an additional defining property: they can spread. In these diseases the tumors may initially be benign, but they often progress to become **malignant tumors** (which are also commonly called **cancers**).

Malignant tumors have two distinguishing features: they can invade neighboring tissues, and the cells can break away and enter the lymphatic system or bloodstream, to be carried to another location where they cross back into tissues to form secondary tumors (Figure 10.1). Spreading to more distant sites in the body is known as **metastasis;** Figure 10.2 shows dissemination via the bloodstream—the cancer cells cross capillary walls and migrate through the extracellular matrix.

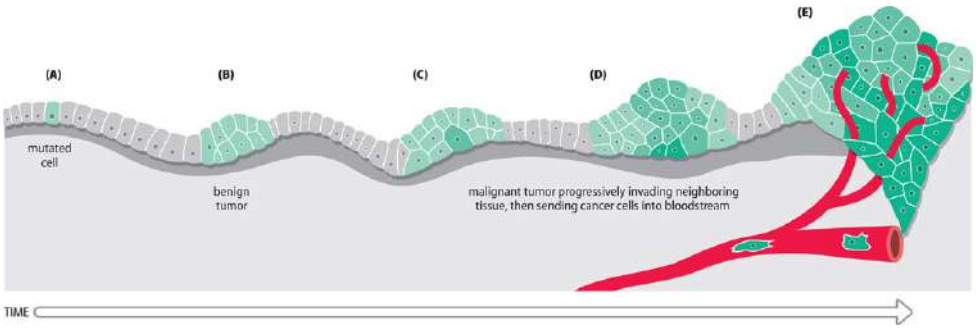


Figure 10.1 Progressive changes in the formation of malignant tumors. The initial mutated cell (A) can develop into a benign tumor (B) through the loss of some normal controls on cell division. Subsequent DNA changes and epigenetic changes can cause tumor cells to lose further

normal controls to become a malignant tumor (C to E) that aggressively invades neighboring tissue. Cells from the malignant tumor can detach themselves and enter the bloodstream (as shown here) or the lymphatic system. In this way they are carried to remote sites in the body where they can exit the circulation and invade neighboring tissues to establish secondary tumors *(metastasis*—for detail of the mechanism see Figure 10.2). (From the website of the National Cancer Institute [http://www.cancer.gov].)

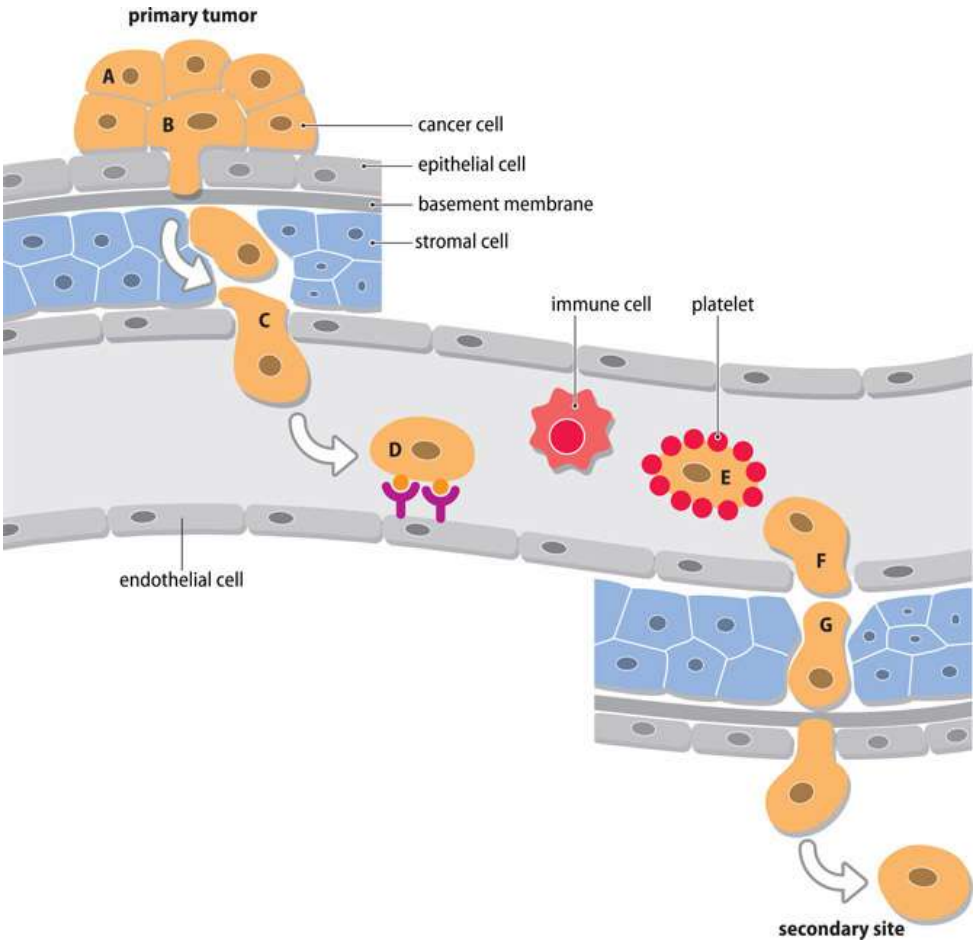


Figure 10.2 The multiple steps taken by metastatic cells to seed secondary tumors. Metastatic cells must first break free from the primary tumor. To accomplish this, cancer cells (A) reduce adhesion to neighboring cells and (B) clear a path for migration into the vasculature-rich **stroma** (connective tissue plus blood vessels). Once at the vasculature, cells can freely enter the bloodstream if the vasculature is discontinuous, as in certain regions of the liver, bone marrow, and kidneys. *Intravasation* (C) is required if the vasculature is continuous; metastatic cells either cause endothelial cells to retract (by releasing compounds such as vascular endothelial growth factor) or induce endothelial cell death (by releasing reactive oxygen species and factors including matrix metalloproteinases). In the bloodstream, cancer cell distribution is

determined by blood flow and interactions between cancer cells and the secondary organs that they colonize: cells can get trapped in narrow capillary beds, such as those of the lung and liver, and can also express receptors that bind to metastasis-supporting sites (D) or to platelets (E), which protect the cancer cells from the immune system. Cancer cells can circulate for more than two hours, suggesting that they do not always become lodged in the first capillary beds that they reach. After reaching the secondary site, cancer cells can leave the bloodstream (F) by *extravasation* (inducing endothelial cell retraction or death). To proliferate in the secondary site, cancer cells co-opt the local environment by releasing proinflammatory compounds and proteinases that induce their neighbors to release growth factors (G). (Adapted from Schroeder A et al. [2012] *Nature Rev Cancer* 12:39–50; PMID 22193407. With permission from Macmillan Publishers Ltd.)

The tumors (also called *neoplasms)* in cancer can be broadly classified as solid or liquid. Solid tumors form discrete masses composed of epithelial or mesenchymal (stromal) cells. "Liquid tumors" are made up of neoplastic cells whose precursors are normally mobile blood cells; they include leukemias and also lymphomas (which, although generally forming solid masses in lymph nodes, are able to travel through the lymphatic system). According to the type of tissues or cells in which they arise, the tumors are classified into different categories (Table 10.1).

TABLE 10.1 MAJOR CATEGORIES OF TUMORS ACCORDING TO TISSUE OR CELLS OF ORIGIN

Tissue/cells of origin	Tumors
Epithelial tissue (single-layer or bilayer)	adenoma (benign); adenocarcinoma (malignant)
Epithelial tissue (multi-layer, as in skin and bladder)	papilloma (benign); squamous cell carcinoma (malignant, in skin); transitional cell carcinoma (malignant, in bladder)

Tissue/cells of origin	Tumors
Blood forming tissue (notably bone marrow)	lymphoma (of lymphocytes); leukemia (of leukocytes)
Stromal (mesenchymal) tissue	benign tumors have the simple -oma suffix; malignant tumors end in -sarcoma. Examples are: fibroma and fibrosarcoma (fibroblasts); osteoma and osteosarcoma (bone); chondroma and chondrosarcoma (cartilage); hemangioma and hemangiosarcoma (endothelial cells)
Glial cells	gliomas

Cancers form when cell division is somehow affected to cause uncontrolled cell proliferation. Changes at the DNA and chromatin levels are the primary contributors. A small minority of human cancers are associated with a specific virus, as described below. Mostly, however, human cancers form because of a series of mutations and epigenetic dysregulation in certain cancer-susceptibility genes.

Recall from Section 4.1 that mutations frequently arise through errors in DNA replication and DNA repair; tissues that have actively dividing cells are therefore prone to forming tumors. Cells divide in development as an organism grows, and some childhood tumors can arise from mutations in cells of the embryo. Although the great majority of our cells are not actively dividing, an adult human has roughly one trillion (10^{12}) rapidly multiplying cells. There is a need to replace certain types of cell that have a high turnover, notably cells in the blood, skin, and the gastrointestinal tract. Thus, for example, each day about 4 % of the keratinocytes in our skin and a remarkable 15–20 % of the epithelial cells of the colon die and are replaced.

The short-lived cells that need to be replaced regularly are ones that interface, directly or indirectly, with the environment, and continuous turnover of these cells is a protective measure. Stem cells are the key cells responsible for manufacturing new body cells to replace the lost cells. As we explain below, considerable evidence suggests that cancers are often diseases of stem cells.

Why cancers are different from other diseases: the contest between natural selection operating at the level of the cell and the level of the organism

We are accustomed to thinking of how Darwinian natural selection works at the level of the organism: the key parameter is the reproductive success rates of individuals. *Selection pressure* is the effect of natural selection on allele frequencies. It ensures that a deleterious allele—one that reduces reproductive fitness—will normally be at a low frequency in the population (according to the penetrance, the frequency will be maintained by new mutation, or by transmission by unaffected carriers). Germline mutations make the key contribution to noncancerous genetic disorders; somatic mutations normally have minor roles.

Cancers are different. Yes, occasionally cancers can run in families, and germline mutations are clearly important in some cases. However, all cancers have multiple somatic mutations and the genetic contribution to cancers is dominated by somatic (post-zygotic) mutation. That happens because natural selection also operates at the level of *cells and cancers show abnormal cell proliferation*.

The balance between cell proliferation and cell death

The principal defining feature of cancer is uncontrolled growth in cell number. Growth occurs when the net balance between cell proliferation and cell death is positive. Cell proliferation is required for growth, but a complicated series of controls ensures that normally our cells do not divide in an uncontrolled fashion; sometimes there is a need for brakes to be applied, and cells are ordered to undergo cell cycle arrest. In the opposite direction is cell death, a natural way of removing inefficient cells, cells that are unwanted, and potentially dangerous cells. Like cell proliferation, cell death can be ordered to occur, and it too is highly regulated.

The mechanisms regulating cell proliferation and cell death involve sophisticated intercellular signaling. Some signaling pathways send instructions for certain cells to proliferate or undergo cell cycle arrest; other pathways induce the death of undesirable cells in some way (programmed cell death, or apoptosis). Classical cancer-susceptibility genes were identified as working to regulate cell division or having direct roles in growth-signaling pathways. Additional cancer-

susceptibility genes were found to have roles in apoptosis, but as well as these types of gene, there are many types of non-classical cancer-susceptibility gene that do not function directly in these areas. Instead, they have indirect roles, functioning in a variety of areas such as DNA repair/genome maintenance, cell metabolism and epigenetic regulation and so on. When such genes are faulty, the resulting increased mutation or epigenetic dysregulation can have consequences for genes directly regulating cell growth or apoptosis. See **Figure 10.3** for a summary.

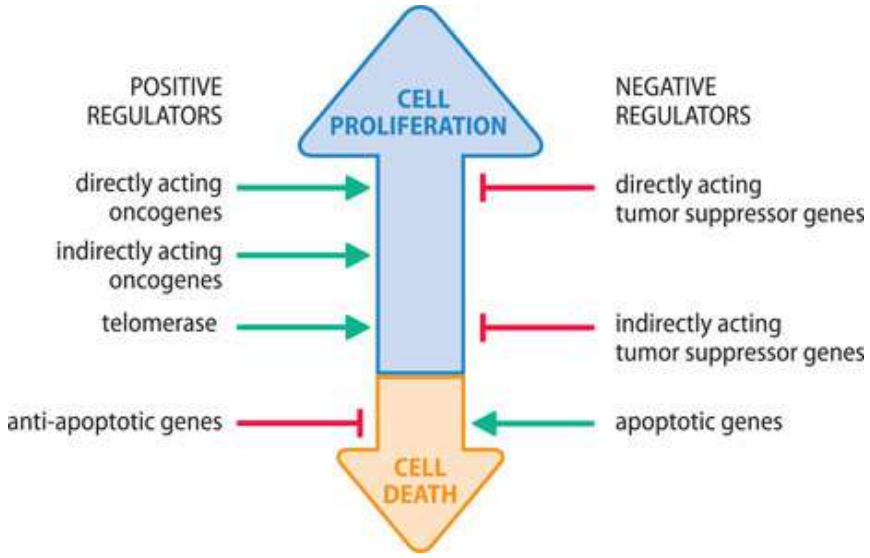


Figure 10.3 Major classes of cancer gene as positive or negative regulators of net cell proliferation (cell growth). Green arrows indicate stimulatory effects on cell proliferation or apoptosis; red T-bars indicate inhibitory effects. Some oncogenes are directly involved in promoting cell division and cell proliferation (by regulating the cell cycle or cell growth signaling pathways). Many other genes have similar effects, but act indirectly, such as oncogenes working in cell metabolism, telomerase and anti-apoptotic genes. Similarly, some tumor suppressor genes control cell proliferation directly; others work indirectly by, for example, regulating genome maintenance or epigenetic pathways. Apoptosis-promoting oncogenes negatively affect net cell proliferation. Note: some tumor suppressor genes, notably TP53, both suppress cell proliferation and promote apoptosis.

Throughout development up to the age of maturity there is an overarching priority for increased numbers of cells to sustain rapid growth of the organism—not in an unconstrained way, of course, but executed according to detailed prescribed body plans and the requirements of intricate tissue architecture, and so

on. But cells are also lost during development. In addition to short-lived cells, many cells are intentionally deleted during development as part of the natural process of sculpting our tissues and organs, and to ensure healthy immune and nervous systems. To distinguish self from nonself, immune system cells with receptors that bind to self-antigens must be deleted, and during nervous system development neurons with unproductive interneuron connections are deleted.

But when we reach adulthood, growth is restricted. Although most of our cells are non-dividing cells by then, a significant minority continue to divide to replace certain types of cell that turn over rapidly, such as skin, blood, and intestinal epithelial cells. Apoptosis is also used in adults to ensure the destruction of both damaged cells (arising as a result of natural wear and tear, or through injury) and potentially harmful cells (such as virus-infected cells).

Why we do not all succumb to cancer?

Cancer is initiated when cells develop capabilities to escape some normal controls limiting cell proliferation, or inducing apoptosis. It is no accident that the cells most likely to give rise to tumors are cells that already possess some elements of the required capabilities. Primarily, they are stem cells (which have either a high intrinsic proliferative capacity, or can be induced by inflammation, tissue damage, and so on to proliferate rapidly), and to a lesser extent populations of cells in the embryo or fetus that transiently undergo rapid proliferation before differentiating.

Cells can escape from these normal controls as a result of mutation in certain control genes. This is where natural selection at the cellular level is important: each successive mutation that disrupts normal controls on cellular proliferation or apoptosis confers an additional selective growth advantage on its descendants. The resulting expansion in mutant cells provides a greater target for successive cancer mutations. As a result, there is strong selection pressure on cells to evolve through a series of stages into tumor cells.

If there is such strong selection pressure on cells to evolve into tumor cells, why do we not all succumb to cancer? Certainly, if we were to live long enough, cancer would be an inevitable consequence of random mutations. However, an opposing force of natural selection works at the level of the organism (to keep us healthy and free from tumors—at least until we have produced and raised children). It involves different mechanisms, not least immunosurveillance to detect and kill

cancer cells, using cytotoxic T cells and natural killer (NK) cells. (Individuals whose immune systems are suppressed are more susceptible to cancer.)

There is, however, an imbalance between natural selection working at the level of the organism and at the level of cells. Luckily for us, natural selection operates over a much longer timescale than does the selection pressure in favor of tumor cell formation. Cancer cells can successfully proliferate and form tumors within an individual person, *but they do not leave progeny beyond the life of their human host*. That is, tumorigenesis processes must start afresh in a new individual. But at the level of the organism, natural selection continues down through generations. Individuals who have efficient cancer defense mechanisms are able to pass on good anti-cancer defense genes to their offspring, and anti-cancer defense systems continue to evolve from generation to generation.

Cancer cells acquire several distinguishing biological characteristics during their evolution

As described in the next section, the development of tumors occurs as a series of stages during which both genetic and epigenetic changes progressively accumulate in cells. During these stages the cells progressively acquire different biological capabilities that mark them out as cancer cells.

By definition, cancer cells show unregulated cell proliferation, and tumors develop by breaking away from normal control systems. They switch off various brakes that normally place limits on cell proliferation and genome instability, and counteract death (apoptosis) signals from neighboring cells. Cancer cells also lose the contact inhibition of normal cells that places limits on cell growth. Partly by overcoming these negative signals, they become masters of their own growth, and go on to acquire the characteristic ability to replicate indefinitely (**Box 10.1**).

To support continued cell proliferation, cancer cells re-adjust their metabolism. Thus, they show increased flux through the pentose phosphate pathway (PPP) and elevated rates of lipid biosynthesis, and they take up and use glucose at much higher rates than normal cells. (The last characteristic can be used by imaging systems to differentiate cancer cells from normal cells, so that the spread of cancer cells in the body can be visualized.)

Although apparently exposed to aerobic conditions, cancer cells nevertheless derive their energy from glycolysis, rather than from oxidative phosphorylation. Glycolysis is normally used by cells in anaerobic conditions; the process involves

converting glucose to first pyruvate, and then lactate, and energy production is inefficient (2 molecules of ATP generated per molecule of glucose, whereas under aerobic conditions, normal cells convert glucose to pyruvate and then pyruvate is catabolized in the tricarboxylic acid cycle, generating up to 36 molecules of ATP per molecule of glucose).

Why, under aerobic conditions, cancer cells normally use the much more inefficient glycolysis system of producing energy (the Warburg effect) remains poorly understood. The switch to glycolysis occurs early in oncogenesis and may be activated to drive cell survival. Interested readers can find a recent review at PMID 33347611.

BOX 10.1 TELOMERE SHORTENING AND SELECTION PRESSURE ON CANCER CELLS TO BECOME IMMORTAL BY ACTIVATING TELOMERASE EXPRESSION

Normal human cells can be grown in culture for limited periods. Fetal cells, for example, can divide 40–60 times in culture before reaching a state of *senescence* after which they cannot grow any further. Cancer cells, however, have unusual growth properties. In culture they do not exhibit contact inhibition, nor require adhesion to a solid substrate, Notably, they can replicate indefinitely, and so are immortal. (The HeLa cell line is the most outstanding human example; developed from a cervical cancer biopsy in the early 1950s, it has been extensively propagated to become the most intensively studied human cell line.)

THE END-REPLICATION PROBLEM

The above observations on the replicative behavior of cells relate to the end-replication problem: how can the extreme ends of linear chromosomes be replicated when new DNA strands grow in the 5¢® 3¢ direction only? During DNA replication, new DNA synthesis is catalyzed by DNA polymerases that use an existing DNA strand as a template. As the replication fork advances, one new DNA strand is made in the same direction as the direction of travel for the replication fork, and can be synthesized continuously in the 5¢ ® 3¢ direction. However, the other strand can be made only by synthesizing successive short pieces of DNA (Okazaki fragments) in the opposite direction to that taken by the

replication fork, and there is a problem with completing the synthesis at the very end (Figure 1). Because of the end-replication problem, using just DNA-dependent DNA polymerases means that a small amount of DNA will be lost from each telomere after every cell division.

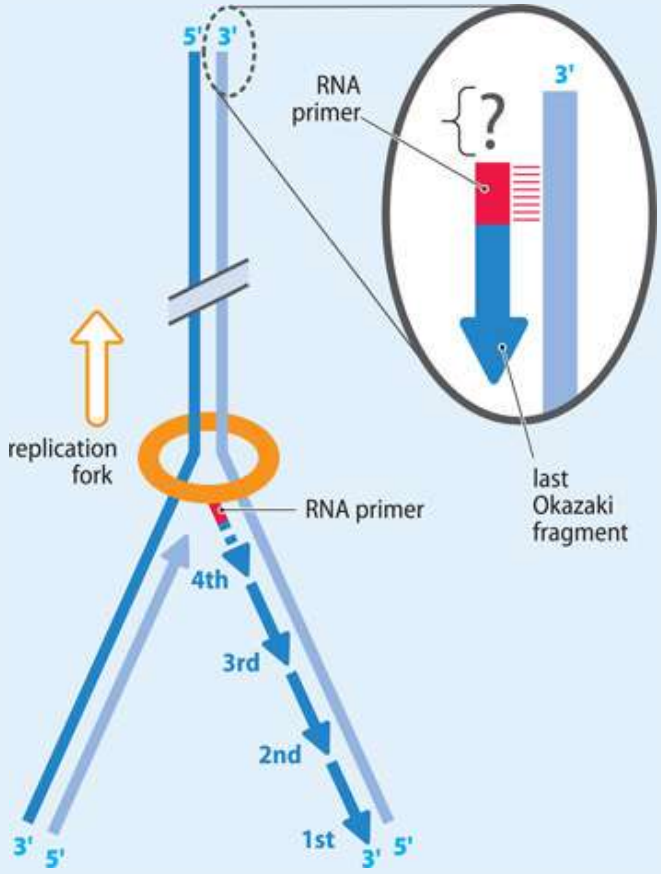


Figure 1 The problem with replicating the extreme ends of DNA in linear chromosomes. In normal DNA replication by DNA-dependent DNA polymerases, an existing DNA strand is used as a template for making a complementary new DNA strand. Here, as the replication fork advances in the upward direction it can synthesize a continuous DNA strand upward in the 5¢® 3¢ direction from one original DNA strand (colored deep blue) but for the pale blue original strand the 5¢® 3¢ direction for DNA synthesis is in the opposite direction to the upward direction of the replication fork. The DNA must be synthesized in short pieces, called Okazaki fragments, starting from a position beyond the last fragment and moving backward toward it. (DNA-dependent polymerases use short RNA primers to initiate the synthesis of DNA, but the RNA primers are degraded, DNA synthesis fills in, and adjacent Okazaki fragments are ligated.) The question mark indicates a problem that is reached at the very end: how is synthesis to be completed when there can be no DNA template beyond the 3¢ terminus?

The telomeres of our chromosomes have tandem TTAGGG repeats extending over several kilobases of DNA (see Figure 1.9 on page 10), but because of the end-replication problem (plus oxidative damage and other end-processing events), the arrays of telomeric TTAGGG repeats normally shorten with each cell division (the number of telomere repeats lost varies between different cell types but is often in the range of 5–20 repeats). When a few telomeres become critically shortened, there is a growth arrest state, at which time DNA damage signaling and cellular senescence is triggered. In the absence of other changes, cells can remain in a senescent state for years.

TELOMERASE SOLVES THE END-REPLICATION PROBLEM

The end-replication problem can be solved—and telomeres restored to full-length—when cells express time DNA damage signaling and cellular senescence is triggered. In the absence of other changes, cells can remain in a senescent state for years. telomerase, an RNA-dependent DNA polymerase. Telomerase is a ribonucleoprotein consisting of a reverse transcriptase and a noncoding RNA (ncRNA). The ncRNA has a hexanucleotide sequence that is complementary in sequence to the telomere repeat; it serves as a template from which the reverse transcriptase can make tandem telomere repeats (**Figure 2**).

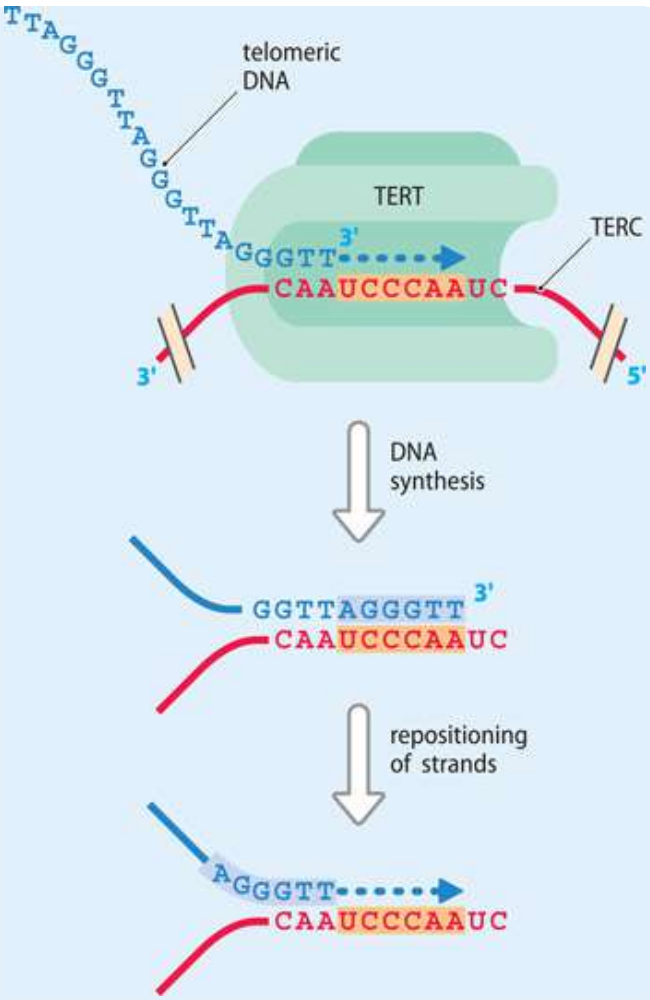


Figure 2 Telomerase uses a reverse transcriptase and a noncoding RNA template to make new telomere DNA repeats. The telomerase reverse transcriptase (TERT) is an RNA-dependent DNA polymerase: it uses an RNA template provided by the other subunit, TERC (telomerase RNA complex). Only a small part of the RNA is used as a template—the hexanucleotide that is shaded—and so the telomeric DNA is extended by one hexanucleotide repeat (blue shading). Repositioning of the telomeric DNA relative to the RNA template allows the synthesis of tandem complementary copies of the hexanucleotide sequence in the RNA template.

Telomerase is expressed during early human embryogenesis (and in embryonic stem cells). However, its expression is subsequently repressed in most somatic cells, but it is active in the male germ line, activated lymphocytes, and stem cells in certain regenerative tissues.

SELECTION PRESSURE ON CANCER CELLS TO ACHIEVE REPLICATIVE IMMORTALITY

The repression of telomerase and the resulting erosion of telomeres in our cells is thought to be yet another defense system to stop cancer from developing during our long lifetimes. Cancer cells require multiple successive mutations to become malignant. After one mutation has led to some growth advantage, about another 20–40 cell divisions might be required to achieve a cell population size that is sufficient for another spontaneous mutation to occur in a cell with the previous mutation. Premalignant cells would therefore often be expected to come up against the barrier of replicative senescence before they have sustained sufficient mutations to form malignant tumors.

Tumor cells are able to bypass replicative senescence by suppressing tumor suppressors, such as p53 and the RB1 retinoblastoma protein. However, after a few additional cell divisions past the point at which senescence normally occurs, the cells enter a *crisis* state. Now the telomeres can be so short that DNA repair mechanisms do not recognize them as legitimate ends of chromosomes; instead, they treat them as double-strand DNA breaks. As a result, chromosomes can undergo end-toend fusions. The resulting chromosomes have two centromeres and may be pulled in opposite directions at mitosis. That causes further broken ends, new cycles of chromosome fusion and breakage, and an acceleration of genome instability.

Rare cells that escape this crisis stage are almost always able to do so by having reactivated expression of telomerase: the telomeres are stabilized and the cell becomes immortal. However, the telomerase produced is not present at excess (the telomeres in cancer cells with stem cell-like properties are generally of the same length or shorter than those in adjacent normal cells). Note: in a small number of cases the ALT (*alternative lengthening of telomeres*) pathway is deployed; here telomeres are lengthened instead via homologous recombination mediated by inactivating mutations in the *ATRX* and *DAXX* genes (which together make a protein complex that deposits histone variant H3.3 into the repetitive heterochromatin of telomeres to promote transcription).

During cancer progression, cancer cells also undergo epigenetic reprogramming so that they can become less differentiated. Solid cancers show a plastic

phenotype, with a differentiated tumor mass and also undifferentiated areas. The latter, notably marking regions that form an invasive front as the cancer spreads, allow flexibility to respond to different environments, and metastases can show striking re-differentiation.

To ensure their survival, cancer cells need to avoid being destroyed by immune system cells, and they develop appropriate counter-attacking measures. Not only that, but they also maximize their ability to survive by invading host tissue and co-opting normal cells to help them, and by sending out cells to form secondary tumors.

It is also common for cancers to gain access to the vascular system by inducing the sprouting of existing blood vessels, whereupon the tumors become linked to the existing vasculature (**angiogenesis**), as in the tumor shown in Figure 10.1E. That then allows tumor cells to escape more readily from a primary tumor and establish secondary tumors, although angiogenesis may often not be necessary for metastasis.

Table 10.2 provides a summary, listing 10 biological characteristics cancer cells acquire that have been proposed as hallmarks of cancer. We expand on some points in Section 10.2, but also provide details on some other of the points in later sections.

TABLE 10.2 TEN ACQUIRED BIOLOGICAL CAPABILITIES PROPOSED AS HALLMARKS OF CANCER BY DOUGLAS HANAHAN AND ROBERT WEINBERG

Acquired biological capability	**Examples of how the biological capability is acquired**
Self-sufficiency in growth signaling	Activate cellular oncogene
Insensitivity to signals suppressing growth	Inactivate *tp53* to avoid p53-mediated cell cycle arrest

IGF, insulin growth factor; TGFβ, transforming growth factor β; VEGF, vascular endothelial growth factor. (Adapted from Hanahan D & Weinberg R [2011] *Cell* 144:646–674; PMID 21376230. With permission from Elsevier.). Note: in addition to these hallmarks, others have been proposed such as epigenetic dysregula-tion, including dedifferentiation (see PMID 33465324 for a recent review).

Acquired biological capability	Examples of how the biological capability is acquired
Ability to avoid apoptosis	Produce IGF survival factor
Replicative immortality	Switch on telomerase (Box 10.1)
Genome instability	Inactivate certain genes involved in DNA repair
Induction of angiogenesis	Produce factor that induces VEGF
Tissue invasion and metastasis	Inactivate e-cadherin
Ability to avoid immune destruction	Paralyse infiltrating cytotoxic T lymphocytes and natural killer cells by secreting TGF or other immunosuppressive factor
Induction of tumor-promoting inflammation	Redirect inflammation-causing immune system cells that infiltrate the tumor so that they help in various tumor functions (see Table 10.3)
Reprogramming energy metabolism	Induce aerobic glycolysis

IGF, insulin growth factor; TGFβ, transforming growth factor β; VEGF, vascular endothelial growth factor. (Adapted from Hanahan D & Weinberg R [2011] *Cell* 144:646–674; PMID 21376230. With permission from Elsevier.). Note: in addition to these hallmarks, others have been proposed such as epigenetic dysregula-tion, including dedifferentiation (see PMID 33465324 for a recent review).

The initiation and multi-stage nature of cancer evolution and why most human cancers develop over many decades

Epidemiology studies have shown that age is a very large factor in cancer incidence (the rate at which it is diagnosed). For example, the age-incidence plots for epithelial cancers suggested that the risk of death from this cause increases roughly as the fifth or sixth power of elapsed lifetime. That observation suggested that perhaps six to seven independent events might be required for an epithelial cancer to develop (if the probability of an outcome is a function of some variable

raised to the power n, a total of n +1 independent events, each occurring randomly, are required for the outcome to be achieved).

The epidemiology studies provided an early indication of the multistage nature of cancer, and a suggestion of the number of critical steps involved. Now we know that as normal cells evolve to become cancer cells, they pick up many somatic changes—both genetic and epigenetic. A small subset of the genetic changes, known as **driver mutations,** result in altered expression of certain key genes (those regulating cell proliferation and apoptosis) so that a growth advantage is conferred on their descendants. Driver mutations are positively selected and causally implicated in cancer development. The remaining mutations are *passenger mutations*.

A cell with a driver mutation that its neighbors lack usually possesses a small growth advantage, of the order of just a 0.4 % increase in the difference between new cell formation and cell death. The growth advantage is small because we have multiple layers of defense against cancer. Many tumor cells succumb to our natural defenses, or are disadvantaged by certain karyotype changes. Despite the high attrition rate of tumor cells, the small growth advantage can ultimately lead to a large mass, containing billions of cells, but that usually takes many years.

Clonal expansion and successive driver mutations

Tumors are monoclonal: all the cells descend from a single starting cell. Strong evidence for that supposition came from studies of B-cell lymphomas. Recall from Section 4.5 that individual B cells in a person make different immunoglobulins, but the cells in individual B-cell lymphomas all make the same type of immunoglobulin.

Preferential clonal expansion of the mutant cells produces an expanded target (more cells) for a second driver mutation to occur in one of the mutant cells. As the process continues, cells progressively acquire more mutations (**Figure 10.4A**) and epigenetic dysregulation, causing them to become ever more like a cancer cell.

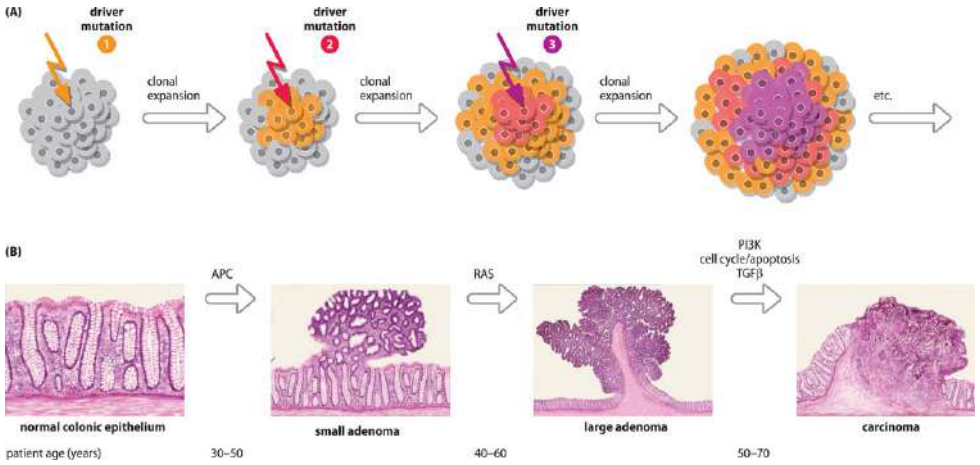


Figure 10.4 Driver mutations in the multistage evolution of cancer. (A) *General process.* Each successive driver mutation gives the cell in which it occurs a growth advantage, so that it forms an expanded clone and thus presents a larger target for the next mutation. Orange cells carry driver mutation 1; red cells have sequential driver mutations 1 and 2; and purple cells have driver mutations 1, 2, and 3. (B) *Genetic alterations and the progression of colorectal cancer.* The major signaling pathways that drive tumorigenesis are shown at the transitions between each tumor stage. One of several driver genes that encode components of these pathways can be altered in any individual tumor. Small and large adenomas appear as intestinal polyps that are benign but can progress to become carcinomas, cancers that invade the underlying tissue. Patient age indicates the time intervals during which the driver genes are usually mutated. Note that this model may not apply to all tumor types. PI3K, phosphoinositide 3-kinase pathway; TGFb, transforming growth factor b pathway. (B, from Vogelstein B et al. [2013] *Science* 339:1546–1558; PMID 23539594. With permission from the AAAS.)

In some cases, just a few driver mutations are required. Figure 10.4B illustrates a classic example: the gradual transformation from normal epithelium to carcinoma in the development of colon cancer. The initial driver mutation is almost always one that affects the Wnt signaling pathway usually through loss of function of the *APC* gene at 5q21, but there can be more flexibility in the order of the subsequent genetic changes.

Cancer develops by accelerating mutation in two major ways

The average rate of mutation in human cells is low (about 10^{-6} per gene per cell) and the majority of cancer-causing mutations are recessive at the cellular level—

both alleles need to be mutated. Cancer might therefore be expected to be highly improbable: the chance that any cell would receive successive mutations, often in both alleles, at several cancer-susceptibility loci would normally be vanishingly small.

Cancer nevertheless is common, and altered expression at a few cancer-susceptibility loci can be sufficient. Cancer is common largely because driver mutations greatly increase the probability of later mutations and epigenetic changes. They do this in two major ways, as listed below.

- *Conferring a growth advantage on cells*. If cells with a driver mutation have an increased growth rate, they will produce more progeny than other cells. Simply by producing an expanded target of mutant cells the probability of a subsequent mutation is increased (see Figure 10.4A).
- *Destabilizing the genome*. This increases the likelihood of later mutations in cancer. Chromosome instability is a feature of most tumor cells, producing grossly abnormal karyotypes with abnormal numbers of chromosomes and frequent structural arrangements that can activate oncogenes or cause a loss of tumor suppressor genes. In some cancers, a form of global DNA instability occurs: mutations in key DNA repair genes result in greatly elevated mutation rates. Increased overall mutations may mean, too, that genes regulating epigenetic modifications are also affected, resulting in altered expression at these types of cancer-susceptibility loci.

Additionally, some types of epigenetic change cause genome instability. In Section 10.3 we explore the detail of genome instability and epigenetic dysregulation, and how they interact in cancer.

The generally late age of cancer onset

Tumors gradually acquire mutations to evolve from benign to malignant lesions. Because that takes some time, cancer is primarily a disease of aging. In self-renewing tissues—such as epithelial cells lining the gastrointestinal tract and genitourinary epithelium—the cells contain DNA that has progressively accumulated mutations through multiple DNA replication cycles in progenitor cells (recall that errors in DNA replication and post-replicative DNA repair are

frequent causes of mutations). Thus a colorectal tumor in a person in their 80s or 90s will have nearly twice as many somatic mutations, mostly inconsequential, as in a morphologically identical colorectal tumor in a person half their age. The difference in ages when the same type of tumor presents will reflect when crucially important driver mutations occurred.

Cells in tissues associated with some other cancers do not replicate, and the tumors associated with these cells have fewer mutations, such as in glioblastomas (advanced brain tumors formed from nonreplicating glial cells) and pancreatic cancers (epithelial cells of the pancreatic duct also do not replicate). Initiating driver mutations in these cases must occur in cells that have had comparatively few mutations.

Childhood cancers and cancers arising in embryonic cells

Some types of cancer commonly arise in childhood. Pediatric tumors often occur in tissues that do not self-renew, and such tumors typically have fewer mutations than adult tumors. However, leukemias and lymphomas, which are diseases of self-renewing blood cells, can also often develop early in life. Here the precursor cells are already mobile and invasive and are thought to require fewer DNA changes than in solid tumors, in which the tumor cells require additional mutations to confer these biological capabilities.

Some childhood cancers—including retinoblastoma, medulloblastoma, nephroblastoma, and Wilms tumor—can arise from an initiating mutation that arises in embryonic cells. Progenitor cells in the embryo resemble cancer cells—they are poorly differentiated and rapidly dividing. If they receive a cancer-predisposing mutation, they are much more likely to develop into tumors at an early stage than more differentiated cells with the same mutation.

Intratumor heterogeneity arises through cell infiltration, clonal evolution, and differentiation of cancer stem cells

Although tumors are considered to be monoclonal (composed of cells derived from a single ancestral cell), that does not mean that the cells in a tumor are the same. Instead, tumors often resemble organs, having quite complicated tissue architectures and being composed of functionally different cells.

A first level of intratumor heterogeneity exists because tumors are usually made up of both tumor cells proper (that originate from a single cell, and so are monoclonal), and also various unrelated cells that infiltrate the tumor from the surrounding environment. Different types of stromal cells, including immune infiltrating cells, become part of the tumor microenvironment and can be redirected to support tumor activities (**Figure 10.5A** and **Table 10.3**).

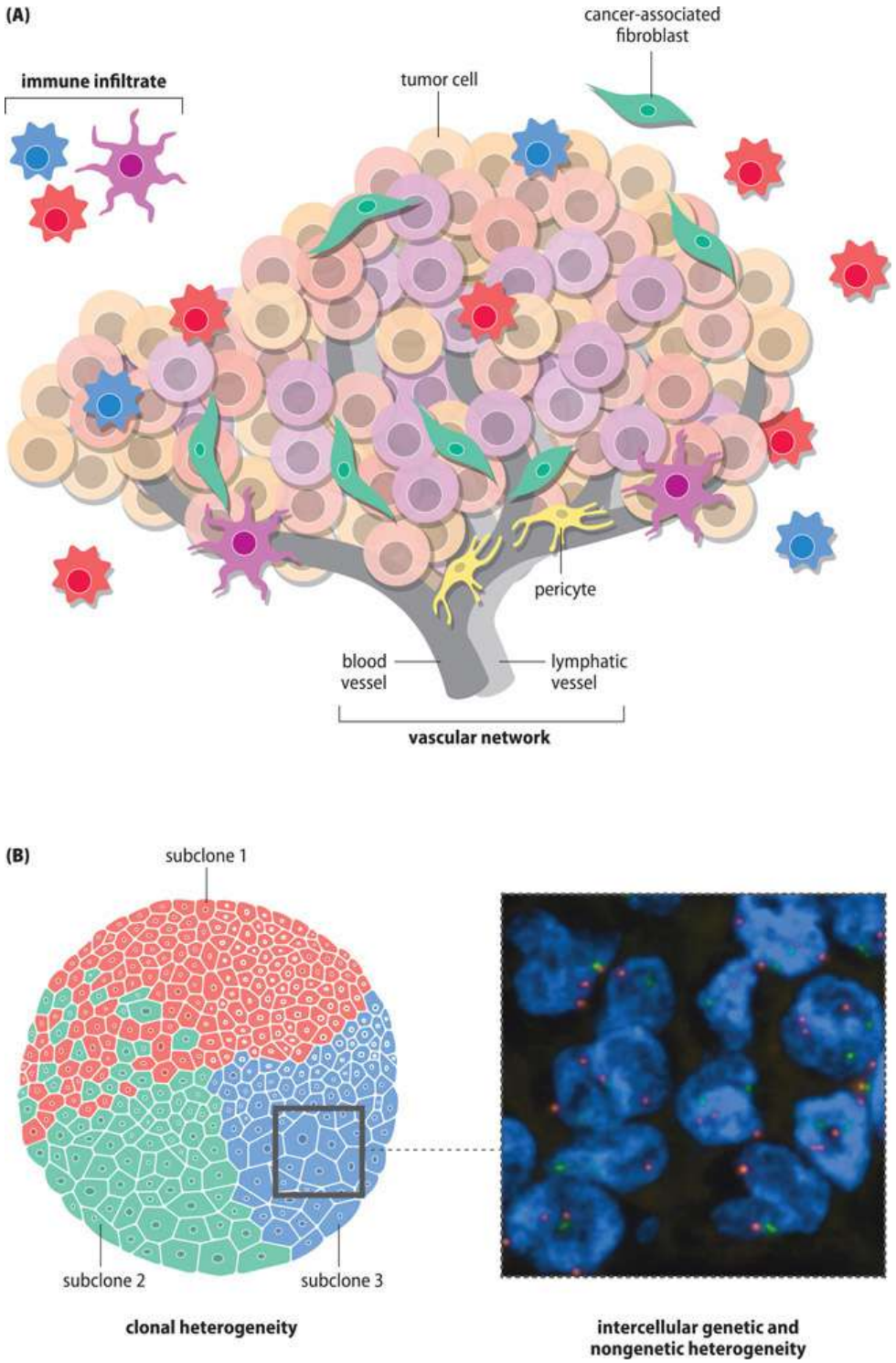


Figure 10.5 Cell heterogeneity within tumors. (A) Tumors as organs. Tumor formation involves co-evolution of neoplastic and non-neoplastic cells in a supportive and dynamic microenvironment that includes different stromal cells—cancer-associated fibroblasts, vascular endothelial cells (including pericytes), and diverse infiltrating immune cells—and the

extracellular matrix. The tumor microenvironment offers structural support, access to growth factors, vascular supply, and immune cell interactions. The immune cells include cell types normally associated with tumor-killing abilities as well as immune cells that can have tumor-promoting properties (see Table 10.3). For an account of the properties of support cells in tumors, see PMID 22439926. (B) Tumor subclones. Different tumor subclones may show differential gene expression due to both genetic and epigenetic heterogeneity. Cells from some subclones may intermingle (subclones 1 and 2) or be spatially separated (subclone 3), sometimes by physical barriers such as blood vessels. Within a subclonal population of tumor cells there may be intercellular genetic and nongenetic variation. For example, in the expanded square (which represents a section taken from a spatially separated subclone), differences in chromosome copy number between cells are revealed by the hybridization signals obtained with fluorescent probes for the centromeres of chromosome **2** (red) and chromosome 18 (green), against a background stain of blue for DNA. (A, From Junttila MR & de Sauvage FJ [2013] *Nature* 501:346–354; PMID 24048067. With permission from Macmillan Publishers Ltd. B, From Burrell RA et al. [2013] *Nature* 501:338–345; PMID 24048066. With permission from Macmillan Publishers Ltd.)

A second level of intratumor heterogeneity exists because the tumor cells proper within a tumor can become functionally distinct from each other. That can happen as a result of differential genetic changes. In addition, differential epigenetic changes can occur and some tumors can clearly be seen to have cells at different stages of differentiation.

Cells that have descended from the originating cancer cell can acquire new mutations conferring some additional growth advantage or other advantageous tumor-associated biological capability. A cell with an advantageous mutation such as this can form a subclone that competes with and outgrows the other cells. The process continues with new subclones competing against previous subclones. Subclones may intermingle or they can be spatially distinct (Figure 10.5B).

After the appearance of successive subclones, a tumor might be dominated by cells from a recent particularly successful subclone (a *clone sweep*) but still contain some cells from previously dominant subclones. Clonal evolution by acquisition of new mutations—both driver and passenger mutations—might therefore explain how functionally different types of tumor cell could arise within the same tumor. Despite being functionally divergent, the cells in the different subclones may or may not be recognizably different in appearance.

Subclones of a primary tumor can also undergo mutations that will drive genetic divergence leading to metastases. A paper published by Wu et al. in 2012 (PMID 22343890) gives the example of clonal selection driving genetic divergence of metastases in medulloblastoma.

In addition to clonal evolution, the concept of cancer stem cells has been invoked to explain intratumor heterogeneity. That is, self-renewing tumor cells have been proposed to give rise to all the different types of tumor cell within a tumor by progressive differentiation (Box 10.2). Because cancer stem cells are very long-lived and can potentially regenerate tumors or seed metastases starting from a single cell, there are important implications for cancer therapy.

Although the concept of cancer stem cells and clonal evolution might seem to provide alternative explanations for intratumor cell heterogeneity, they are not mutually exclusive. There is some evidence that even the stem cells within some tumors are heterogeneous as a result of mutation-induced divergence.

TABLE 10.3 DIFFERENT CATEGORIES OF STROMAL CELL TYPES CAN SUPPORT THE TUMOR MICROENVIRONMENT

Stromal cell category	Examples of cell types	Functions in support of:			
		Mitogenesis	Angiogenesis	Tissue invasion	Metastasis
Infiltrating immune cells	macrophages	+	+	+	+
	mast cells	+	+	+	
	neutrophils	+	+	+	+
	T cells (notably of the Th2-CD4 class and regulatory T cells); B cells	+			

The table gives a quite selective list, both of the different stromal cell types that support tumors and of theirfunctions. For a fuller account, see Table 2 of Hanahan D & Coussens LM (2012) *Cancer Cell* 21:309–322; PMID 22439926.

Stromal cell category	Examples of cell types	Functions in support of:			
		Mitogenesis	Angiogenesis	Tissue invasion	Metastasis
Cancer-associated fibroblastic cells	activated tissue fibroblasts	+	+	+	+
Endothelial cells	endothelial tip, stalk, tube		+		
Pericytes	mature/immature pericytes		+		

The table gives a quite selective list, both of the different stromal cell types that support tumors and of theirfunctions. For a fuller account, see Table 2 of Hanahan D & Coussens LM (2012) *Cancer Cell* 21:309–322; PMID 22439926.

BOX 10.2 CANCER AS A DISEASE OF STEM CELLS

Somatic cells may have rather short lives—only about a week or so, on average, in the case of intestinal epithelial cells. Short-lived cells need to be replaced periodically by new cells produced ultimately from the relevant tissue stem cells. The stem cells are capable of two types of cell division: symmetrical cell division and asymmetrical cell division. If the numbers of stem cells get too low for any reason, they can quickly regenerate by multiple successive symmetrical cell divisions, each producing daughter cells identical to the parent stem cell.

Asymmetrical stem cell divisions are reserved for making differentiated cells. In this case, when the stem cell divides it gives rise to one daughter cell that is identical to the parent cell (step a in **Figure 1**), plus a more differentiated, **transit-amplifying cell** (step b in Figure 1).

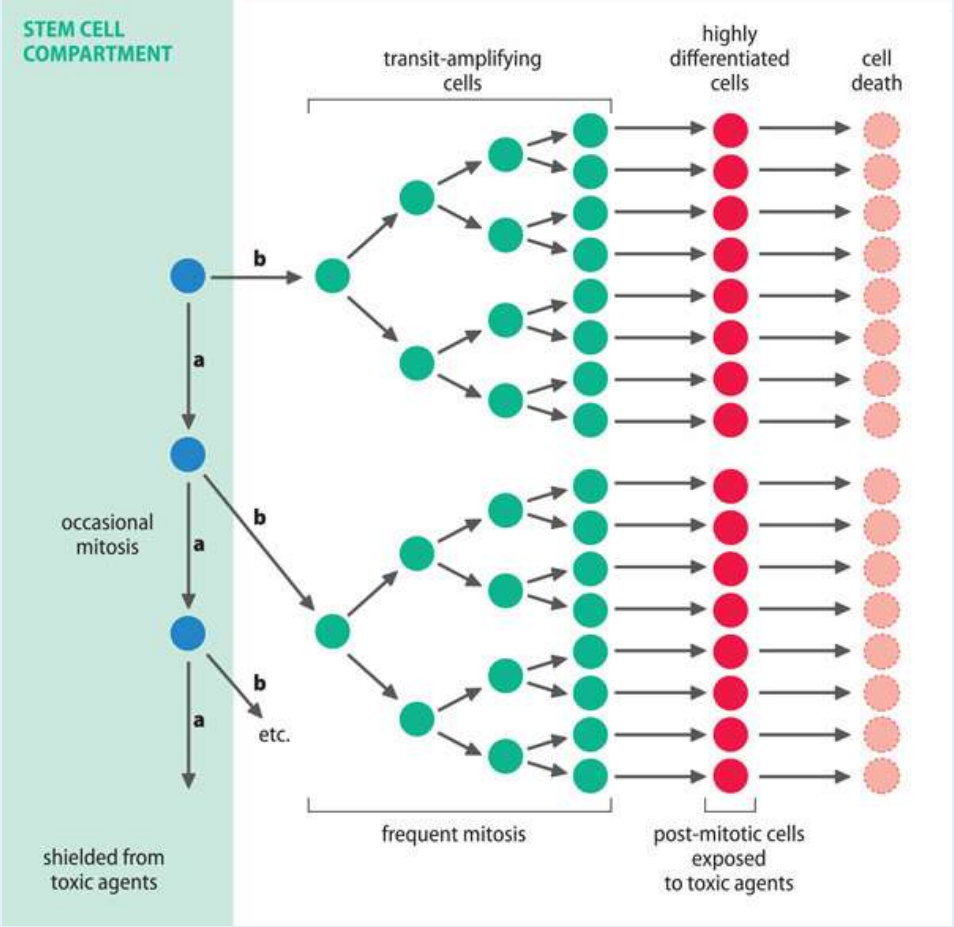


Figure 1 Epithelial tissue as an example of cell differentiation from stem cells and protection of the stem cell genome. Each stem cell *(blue)* divides only occasionally in an asymmetric fashion (steps a and b) to generate a new stem cell daughter and a more differentiated transit-amplifying daughter cell *(green*). Transit-amplifying cells undergo repeated rounds of growth and division, leading to exponential increase in cell numbers. Eventually, the products of these cell divisions further differentiate into post-mitotic highly differentiated cells (*red*). The highly differentiated cells are often in direct contact with various toxic agents, and are frequently shed (so that any harmful mutations that arise in these cells are quickly lost from the tissue). The stem cells are protected from the potentially mutagenic effects of toxic agents because they are shielded by an anatomical barrier. (Adapted from Weinberg RA [2014] *Biology of Cancer*, 2nd edn., Garland Science.)

Newly formed transit-amplifying cells go through multiple symmetrical cell divisions to produce very large numbers of cells that subsequently undergo differentiation to give rise to the highly differentiated, comparatively short-lived

cells. Because the latter cells have short lives, mutation in them could never lead to cancer. Instead, cancer must arise in a longer-lived progenitor cell.

Because the lineage of stem cells represents the only stable repository of genetic information within the tissue, the genomes of stem cells need to be protected as far as possible from mutation. First, the stem cell compartment is physically separated to reduce contact with potential mutagens. For intestinal epithelial cells, for example, the stem cell compartment lies at the base of the intestinal crypts (see Figure 8.11B on page 262 for the latter); here they are far removed from the epithelial cell lining that comes into contact with potentially hazardous mutagens in our diet. Secondly, because transit amplifying cells can expand exponentially from a single stem cell, stem cells may often be comparatively quiescent: by rarely needing to divide, mutations arising from DNA replication errors are minimized (but intestinal epithelial stem cells divide quite regularly).

Despite efforts to maintain their genomes, the very long lifetimes of stem cells make them targets for muta-genesis to form cancer stem cells. Additionally, mutations can be conveyed indirectly into the stem cell pool after mutated transit amplifying cells are dedifferentiated by epigenetic changes to become cancer stem cells. By being relatively resistant to cytotoxic chemicals and radiation, and able to keep on regenerating the more differentiated tumor cells, cancer stem cells pose problems for cancer treatment.

There is now substantial evidence for stem cells in many cancers, including in solid tumors. One of the early pieces of supportive evidence came from analyses of many types of leukemia. For example, the Philadelphia chromosome (a specific chromosomal translocation predisposing to chronic myeloid leukemia [CML; see below]) is often seen in different types of blood cells (B and T lymphocytes, neutrophils, granulocytes, megakaryocytes, and so on) in CML patients. The Philadelphia chromosome is presumed to arise in a precursor of all those different blood cell types, a hematopoietic stem cell. The idea of cancer stem cells is also supported by hierarchical cell organizations for certain types of cancer. In some types of neuroblastoma and myeloid leukemia, for example, the cancer evolves so that some tumor cells differentiate, and have limited capacity for proliferation despite retaining the oncogenic mutations of their malignant precursors.

Further evidence came from flow cytometry, which allows separation of phenotypically distinct subpopulations of live cancer cells that can be studied

after transplanting them into immunocompromised mice. Using this approach, it became clear that only a small proportion of cancer cells in leukemia and breast cancer proliferate extensively, and they express specific combinations of cell surface markers. Breast-cancer-initiating cells, for example, are found to be a minority population of $CD44^{+}CD24^{-}$ cells and leukemia initiating cells are a minority population of $CD34^{+}CD38^{-}$ cells. Lineage studies have also supported the existence of cancer stem cells—we describe in Section 10.5 how single-cell genomics is transforming cancer research.

10.2 ONCOGENES AND TUMOR SUPPRESSOR GENES

By 2020, 568 human genes had been identified as mutational cancer driver genes, having a causal role in cancer as a result of mutation. The relevant data were obtained by different approaches: analyzing tumor-associated chromosomal rearrangements (notably translocations), identifying tumor-specific changes in gene copy number, and identifying tumor-specific mutations (after comparing tumor DNA sequences with the corresponding DNA sequence in normal cells from the same individual).

Two fundamental classes of cancer gene

The key cancer-susceptibility genes—those in which driver mutations occur—can be grouped into two fundamental classes, according to how they work in cells. Some are dominant at the cellular level: mutation of a single allele is sufficient to make a major, or significant, contribution to the development of cancer. Others are recessive: both alleles need to be inactivated to make a significant contribution to cancer.

Oncogenes are the exemplars of dominantly acting cancer-susceptibility genes. In our cells the normal copies of these genes (sometimes called protooncogenes) often function in growth signaling pathways to promote cell proliferation or inhibit apoptosis, but as we describe below they can also work in other cellular functions. An activating mutation in a proto-oncogene can result in inappropriate constitutive high-level expression (instead of being switched on just when needed). An

activating mutation like this in just a single allele of a proto-oncogene can make a significant contribution to the tumorigenesis process.

Tumor suppressor genes are the exemplars of recessive cancer-susceptibility genes. Normal copies of classical tumor suppressor genes work in the opposite direction to oncogenes—to suppress cell proliferation (by inducing cell cycle arrest) or to promote apoptosis of deviant cells. When both alleles of a classical tumor suppressor gene are inactivated, that locus can make a significant contribution to cancer development.

A common analogy imagines an oncogene as the accelerator of a car and a tumor suppressor as the brake. The car will run out of control if the accelerator is jammed on (inappropriately activated) or if the brake fails. The cell is more complicated than this analogy allows: it has many different types of accelerator and brake to regulate cell growth and turnover, and usually several of the cell's accelerators and brakes need to be faulty to cause real damage.

As well as standard oncogenes and tumor suppressor genes, various other types of cancer-susceptibility gene have been identified that, when mutated, can assist tumor development. As described below, some work in DNA repair and genome maintenance. Some others support certain biological capabilities of cancer cells; they include, among others, genes encoding telomerase and proteins involved in energy metabolism and angiogenesis.

Viral oncogenes and the natural roles of cellular oncogenes

Oncogenes were discovered after it became clear that certain cancers in chickens and rodents were induced by viruses. (Note that most human cancers are not caused by viruses. Nevertheless, some viruses are implicated in specific human cancers: Epstein–Barr virus in nasopharyngeal carcinoma and lymphomas; some papillomaviruses in cervical and oropharyngeal squamous cell carcinoma; chronic hepatitis B virus infection in hepatocellular carcinoma; acute transforming human T-cell lymphotropic virus in acute T-cell leukemia; and human herpesvirus-8 in Kaposi's sarcoma.)

Among the viruses found to cause animal cancers were types of acute transforming retrovirus (also called oncoretroviruses) that could make cells in culture change their normal growth pattern to resemble that of tumors (a process known as **transformation**). Whereas normal versions of these retroviruses had the three standard transcription units *(gag, pol,* and *env*), the oncoretroviruses had an

altered genome in which part of the viral genome had been replaced by an altered copy of a cellular gene (a *proto-oncogene*). The copy of the cellular proto-oncogene in an oncoretrovirus is located close to powerful viral promoter/enhancer sequences that ensure inappropriate high-level expression that drives abnormal cellular proliferation, leading to cancer.

The normal cellular gene, the *proto-oncogene,* plays a role in growth-signaling pathways (as a growth factor, growth factor receptor, cell cycle regulator, transcription factor or some other protein working in signal transduction), or by inhibiting apoptosis. It normally promotes cell proliferation only when there is a natural need for cell proliferation. But it too becomes an oncogene when inappropriately expressed, as detailed in the following section

How normal cellular proto-oncogenes are activated to become cancer genes

Proto-oncogenes are activated by a DNA change that is dominant at the cellular level (and normally affects just a single allele). In the subsections below, we describe the three ways in which this can occur. Two of the three types of DNA change result in enhanced gene expression, so that the gene involved does not respond to normal inhibitory signals. The third class is made up of activating point mutations that alter how the protein behaves (see **Figure 10.6**).

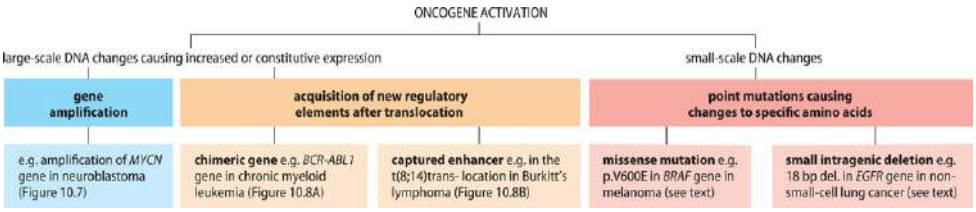


Figure 10.6. Three major ways in which cellular proto-oncogenes are activated to become cancer genes.

Note that cells have multiple anti-cancer defense systems, and the activation of a single cellular proto-oncogene is usually not oncogenic by itself. If we experimentally activate a cellular proto-oncogene in cultured cells, the usual effect is to induce cell cycle arrest (the abnormal proliferative signals usually induce cellular defense mechanisms that shut down cell proliferation); multiple genetic (and epigenetic) changes are needed to induce cancer.

Activation by gene amplification

Tumor cells often contain abnormally large numbers—often hundreds of copies—of a structurally normal oncogene. The *MYCN* oncogene, for example, is frequently amplified in late-stage neuroblasts (**Figure 10.7A**) and in rhabdomyosarcomas; *ERBB2* (also called *HER-2*) is often amplified in breast cancers. The gene amplification mechanism is not simple tandem amplification; instead, there seem to be complex rearrangements that bring together sequences from several different chromosomes. The amplification may manifest itself in two forms:

- *double minutes*, an extrachromosomal form made up of tiny, paired acentric chromatin bodies that are separated from chromosomes and contain multiple copies of just a small set of genes (Figure 10.7B)
- *homogeneously staining regions,* a corresponding intrachromosomal form in which multiple repeated copies integrate into chromosomes.

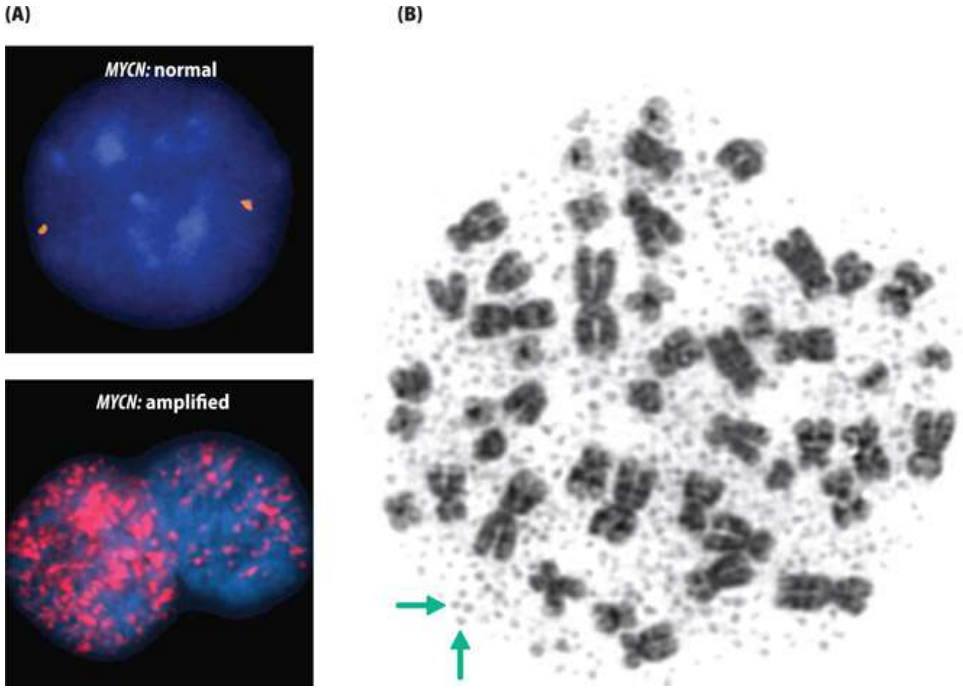


Figure 10.7 Amplification of the *MYCN* gene and formation of double minutes in neuroblastoma cells. (A) Fluorescence *in situ* hybridization (FISH) images using a labeled *MYCN* gene probe, showing two copies of the gene (red signals) in normal cells against a background of DNA staining (shown in blue). In neuroblastoma cells, the *MYCN* gene can undergo extensive amplification to produce many dozens or even hundreds of *MYCN* gene

copies (as shown at the bottom). (B) A metaphase chromosome preparation from a neuroblastoma tumor sample, showing *double minute* chromosomes (which appear as a cloud of very small dots; arrows indicate two of them). (A, Courtesy of Nick Bown, NHS Northern Genetics Service, Newcastle upon Tyne, UK; B, Courtesy of Paul Roberts, NHS Cytogenetics Service, Leeds, UK.)

Translocation-induced oncogene activation

Chromosomal translocations occur when DNA molecules receive double-strand breaks and are then rejoined incorrectly so that pieces of different DNA molecules are joined together. When that happens, an oncogene is often inappropriately transcriptionally activated and so there can be a selective growth advantage.

Translocations that activate oncogenes are common in cancer (more than 300 cancer-associated translocations are listed within the Cancer Gene Census section of the COSMIC database described in Section 10.4). In many cases, the translocations result in the formation of clearly chimeric genes that result in the constitutive expression of oncogene sequences. In other cases, the onco-gene sequence is not interrupted by a breakpoint; instead it is simply brought into close proximity to regulatory sequences in another gene that is actively expressed (see **Table 10.4** for some examples).

TABLE 10.4 EXAMPLES OF ONCOGENE ACTIVATION BY TRANSLOCATION

Tumor type	Oncogene (location)	Interacting gene (location)
Acute lymphoblastoid leukemia (ALL)	*MLL* (11q23)	*AF4* (4q21), *AF9* (9p22) *AFX1* (Xq13), *ENL*(19p13)
Acute myeloid leukemia (AML)	*FUS* (16p11)	*ERG* (21q22)

Note that certain oncogenes such as *MLL* participate in translocations with many other genes and that immunoglobulin genes (such as *IGH*), and T-cell receptor genes (such as *TRD*) are frequently involved in oncogene-activating translocations to cause B- or T-cell cancers. For a complete list, go to the Cosmic database at https://cancer.sanger.ac.uk/census/ then move to the Breakdown list and select Translocations.

* See Figure 10.8A.

Tumor type	Oncogene (location)	Interacting gene (location)
Acute promyelocytic leukemia	*PML* (15q24)	*RARA* (17q21)
Burkitt's lymphoma	*MYC* (8q24)	*IGH* (14q32)
Chronic myeloid leukemia (CML)	*ABL* (9q34)*	*BCR* (22q11)*
Ewing sarcoma	*EWS* (22q12)	*FLI1* (11q24)
Follicular B-cell lymphoma	*BCL2* (18q21)	*IGH* (14q32)
Tcell leukemia	*LMO1* (11p15), *LMO2* (11p13), L1(1p32)	*TRL1* (14q11)

Note that certain oncogenes such as *MLL* participate in translocations with many other genes and that immunoglobulin genes (such as *IGH*), and T-cell receptor genes (such as *TRD*) are frequently involved in oncogene-activating translocations to cause B- or T-cell cancers. For a complete list, go to the Cosmic database at https://cancer.sanger.ac.uk/census/ then move to the Breakdown list and select Translocations.

* See Figure 10.8A.

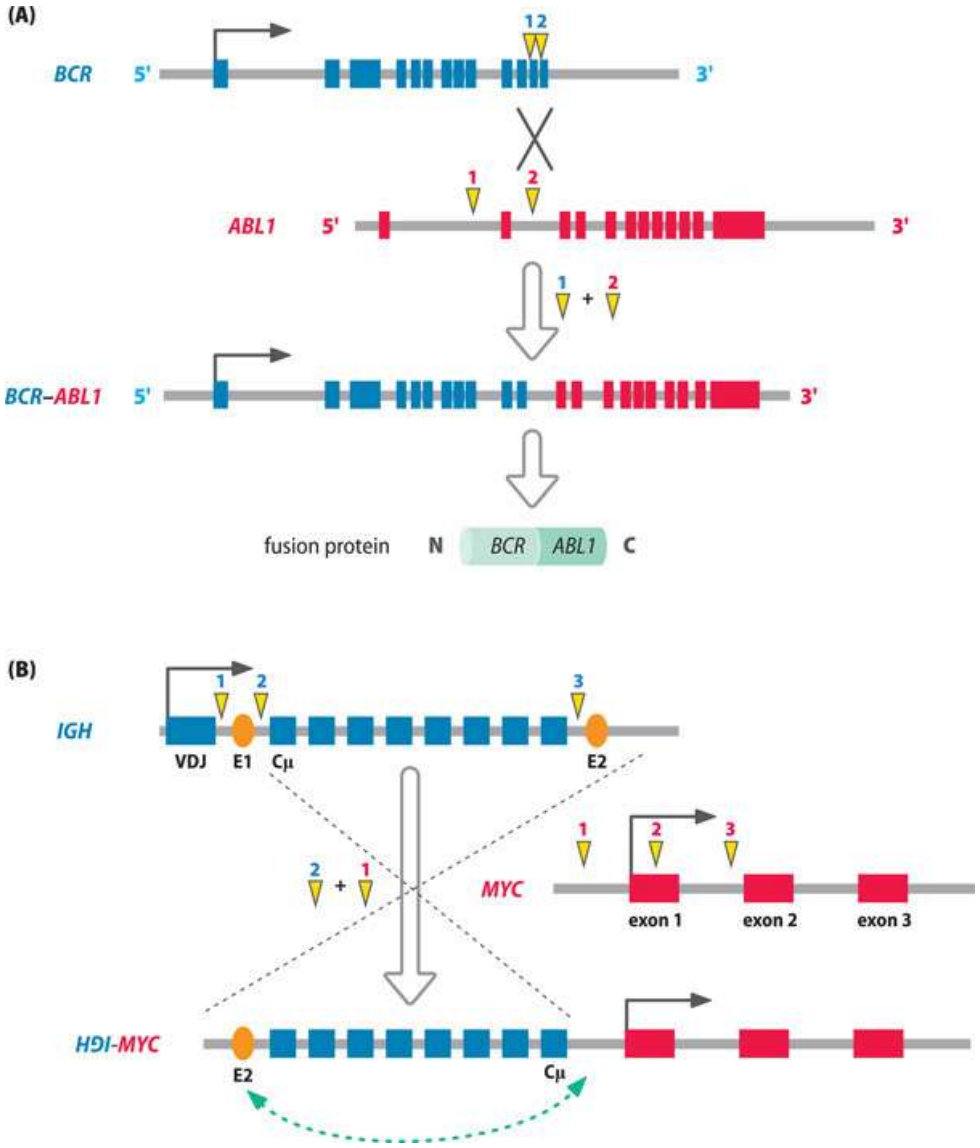


Figure 10.8 Two ways in which translocation can result in oncogene activation by ensuring inappropriate expression. (A) Chimeric gene formation. This shows formation of the chimeric *BCR-ABL1* gene in acute myeloid leukemia, permitting inappropriate expression of the *ABL1* oncogene as a fusion BCR-ABL1 protein expressed through the promoter and regulatory sequences provided by the *BCR* gene. Blue and red vertical bars represent exons. Vertical yellow darts indicate observed breakpoints in patients; here we show recombination at breakpoint 1 in the *BCR* gene, and at breakpoint 2 in the *ABL1* gene. The resulting *BCR–ABL1* chimeric gene produces a large protein with constitutively active tyrosine kinase activity, which does not respond to normal controls. (B) Enhancer capture. The common t(8;14) translocation in Burkitt's lymphoma brings a B-cell specific enhancer (yellow oval) from the *IGH* (immunoglobulin heavy chain) gene into close proximity to the *MYC* gene so that MYC is inappropriately activated in B-cells. Here we show the chromosome that results after the use of breakpoint 2 in the *IGH* gene

and breakpoint 1 in the *MYC* gene. Note that on the translocation chromosome the two sense strands of the genes are on opposing DNA strands so that the 3¢ end of the *IGH* gene is distant from the 5¢ end of the *MYC* gene. The E2 B-cell enhancer is nevertheless close enough to the promoter of the *MYC* gene so that they interact (dashed green arrow), driving strong inappropriate expression of the *MYC* gene in B cells.

The Philadelphia (Ph^1) chromosome, occurring in 90 % of individuals with chronic myeloid leukemia, illustrates how a translocation gives rise to cancer via a chimeric gene. It results from a balanced reciprocal translocation with breakpoints near the start of the *ABL1* oncogene at 9q34 and close to the end of *BCR* gene at 22q11 (**Figure 10.8A**). The resulting *BCR-ABL1* fusion gene on the Philadelphia chromosome (with the *ABL1* coding sequence positioned downstream of the *BCR* gene sequence and *BCR* promoter) produces a large protein that carries the ABL1 polypeptide sequence at its C-terminal end. This fusion protein acts as a growth-stimulating tyrosine kinase that is *constitutively* active and so drives cell proliferation.

Tumors of B and T cells, including various lymphomas and leukemias, often result from translocations with breakpoints in an immunoglobulin heavy-chain or light-chain gene, notably *IGH,* or in a T-cell receptor gene such as *TRA, TRB,* or *TRD*. Recall from Section 4.5 that developing B and T cells are quite exceptional cells because of the requirement for programmed DNA rearrangements that rearrange, respectively, immunoglobulin (Ig) and T-cell receptor (TCR) genes in order to make cell-specific Ig or TCR chains. Because of the natural need to produce double-strand breaks in the Ig or TCR genes, there is a higher chance that these genes will participate in translocations.

Because the large Ig and TCR genes contain many different enhancer sequences, translocations can often result in the transcriptional activation of an oncogene that lies close to the reciprocal breakpoint. As a result, some translocations activate an oncogene simply by bringing a T- or B-cell specific enhancer in close proximity to the oncogene promoter (see Figure 10.8B)

Gain-of-function mutations

Oncogenes can also be activated by certain point mutations that make a specific change at one of a few key codons (often a missense mutation, but sometimes

small deletions of a few codons can change the behavior of the relevant protein). Activating mutations in some cellular oncogenes are particularly common, especially when the genes make a product that links different biological pathways connected to cell proliferation and growth.

Take, for example, the human Ras oncogenes—*HRAS, KRAS*, and *NRAS*—that make highly related 21 kDa Ras proteins with 188 or 189 amino acids. The Ras proteins work as GTPases and mediate growth signaling by receptor tyrosine kinases in mitogen-activated protein (MAP) kinase pathways. Heavily implicated in cancer, they act as signaling hubs (a single Ras protein interacts with multiple intercellular signaling proteins and can transmit a signal from a receptor tyrosine kinase to various downstream signaling pathways, thereby affecting multiple processes). About one in six human cancers has activating mutations in a RAS gene, most commonly in *KRAS* (which is naturally expressed in almost every tissue). More than 99 % of the activating Ras mutations are in one of only three key codons, codon 12 (Gly), codon 13 (Gly), and codon 61 (Gln).

The bias toward missense mutations, and the very narrow distribution of where the mutations occur, distinguishes oncogenes from tumor suppressor genes (see **Figure 10.9** for examples). In some cases, small intragenic deletions that remove a few codons are observed that can also result in a change of function. For example, the c.2240_2257del18 mutation in the *EGRF* gene is commonly found in non-small-cell lung cancer, and replaces a heptapeptide sequence of the EGFR protein by a serine. The mutation affects an ATP-binding pocket and the effect is to enhance signaling, a gain of function.

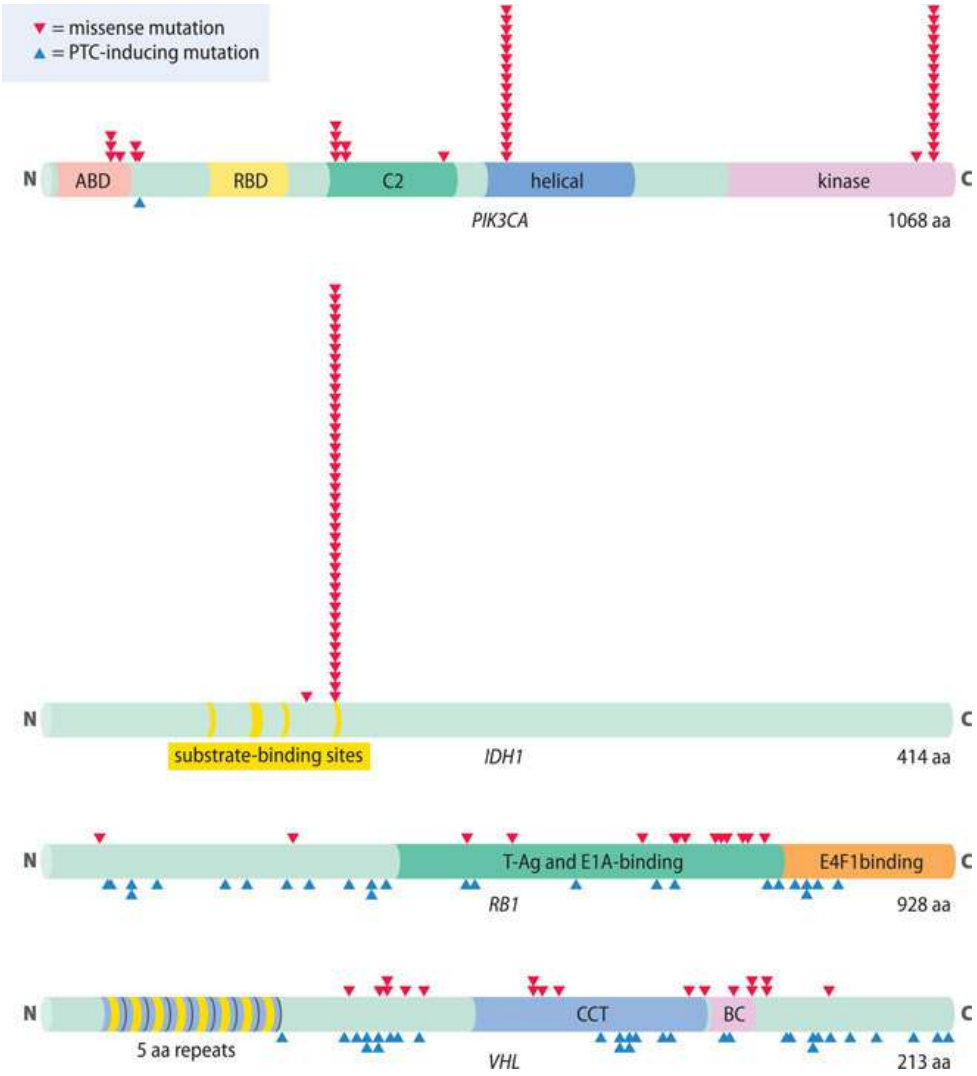


Figure 10.9 Oncogenes differ from tumor suppressor genes in the distribution and range of cancer-associated mutations. The distributions of cancer-associated missense mutations (red arrowheads) and mutations introducing a premature termination codon (PTC; blue arrowheads) are mapped to the corresponding regions of protein products for two representative oncogenes *(PIK3CA* and *IDH1)* and two tumor suppressor genes *(RB1* and *VHL).* Colored bars on the pale green background represent functional domains and motifs. The data were collected from genome-wide studies annotated in the COSMIC database (release version 61). For *PIK3CA* and *IDH1,* mutations obtained from the COSMIC database were randomized, and the first 50 are shown. For *RB1* and *VHL,* all mutations recorded in COSMIC were plotted. Note the predominance of missense mutations in the oncogenes and how they are restricted to just a very few codons. Abbreviation: aa, amino acid residues. (From Vogelstein B et al. [2013] *Science* 339:1546–1558; PMID 23539594. With permission from the AAAS.)

It should be noted that even advanced cancers retain some characteristics of their tissue of origin, and so a gene that might behave as an oncogene in one type of tumor may behave differently in a tumor originating from a different tissue. Thus, for example, the frequent observation of specific missense mutations in the *NOTCH1* gene in lymphomas and leukemias indicate that here *NOTCH1* behaves as an oncogene. But in squamous cell carcinomas, *NOTCH1* mutations are nonrecurrent and usually inactivating, suggesting that in these tumors *NOTCH1* might behave as a tumor suppressor.

Tumor suppressor genes: normal functions, the two-hit paradigm, and loss of heterozygosity in linked markers

Tumor suppressor genes make products that keep cells under control by restraining cell proliferation either directly or indirectly, and have been classified into different groups as listed below.

- *Gatekeeper genes* directly restrain cell proliferation. Their products regulate cell division—by regulating the cell cycle and inducing cell cycle arrest, as required, or by working in upstream growth signaling pathways —or they promote apoptosis.
- *Caretaker genes,* indirectly restrain cell proliferation by helping to maintain the integrity of the genome
- *Landscaper genes* indirectly restrain cell proliferation by controlling the stromal environment in which the cells grow.

Unlike oncogenes, a tumor suppressor gene contributes to cancer when the gene is lost or inactivated in some way. Whereas mutated oncogenes act in a dominant manner at the cellular level, mutated tumor suppressor genes often act in a recessive manner. For classical tumor suppressor genes, inactivation of one copy of a tumor suppressor gene has little effect; the additional loss or inactivation of the second gene is required in the tumorigenesis process. For these genes, the tumor suppressor locus needs to sustain two “hits” to make a signifi-cant contribution to tumorigenesis.

Familial cancers and the two-hit paradigm

From what we have described so far, the idea of familial cancers might seem strange; nevertheless, they do account for a minority of cancers. Familial cancers nearly always involve inheritance of a loss-of-function allele in a tumor suppressor gene (but see below for the example of inherited mutations in the *RB1* oncogene).

The two-hit hypothesis proposed by Alfred Knudson explained why certain tumors can occur in hereditary or sporadic forms. In the hereditary form, one inactivating mutation (the first hit) in a tumor suppressor gene is inherited and the second hit occurs in the somatic cancer progenitor cell; in the sporadic form, two successive inactivating hits, one in each allele, occur in a somatic cell to initiate tumorigenesis (see **Figure 10.10**).

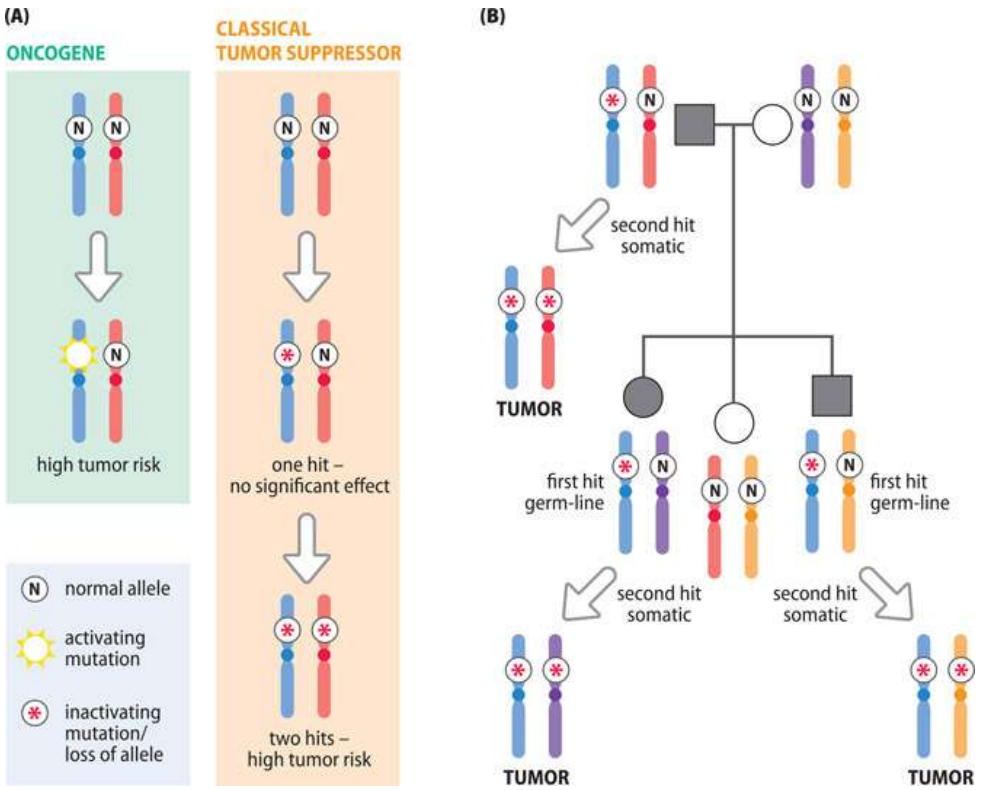


Figure 10.10 Classical tumor suppressor genes and the two-hit hypothesis. (A) Activating mutations in a single allele of an oncogene are sufficient to confer a high risk of tumorigenesis. For a classical tumor suppressor locus to make a significant contribution to tumorigenesis, both alleles need to lose their function (the loss of function may occur through mutational inactivation or loss of the allele, or sometimes epigenetic silencing). Some tumor suppressors do not follow this simple model. (B) Cancers due to mutations at a classical tumor suppressor locus are recessive at the cellular level (both alleles need to be inactivated) but cancer susceptibility can still be dominantly inherited. Inheritance of a single germline mutation (first hit, on the pale blue

chromosome here) means that each cell of the body already has one defective allele and there is a very high chance of some cells receiving a second (somatic) hit. In sporadic forms of the disease, tumors are thought to arise by two sequential somatic mutations in the same cell.

Retinoblastoma, a cancer of the eye that represents 3 % of childhood cancers, provided the first support for the two-hit hypothesis. In retinoblastoma, tumors can occur in both eyes or in one eye. People with bilateral tumors often transmit the disorder to their children, but the children of a person with a unilateral retinoblastoma usually do not have retinoblastoma.

Statistical modeling indicated that hereditary cases of retinoblastoma probably developed after only one somatic mutational event. People with bilateral retinoblastomas were postulated to have inherited an inactivating mutation in one copy of a retinoblastoma-susceptibility locus, now called *RB1*; in that case each nucleated cell in the body would have one inactive *RB1* allele. Retinoblastomas develop from many poorly differentiated retinoblast progenitor cells that proliferate rapidly. There is therefore a high chance that within a population of a million or so retinoblasts carrying an inactivated *RB1* allele, more than one cell sustains an additional inactivating mutation in the second *RB1* allele. If multiple tumors can form, bilateral tumors are likely to occur.

If, however, two normal *RB1* alleles have been inherited, each tumor must occur by two successive hits at the *RB1* locus in one somatic cell. Unless the first somatic mutation just happened to occur very early in embryogenesis, the chances that two sequential somatic mutations would cause a loss of function of both *RB1* copies in more than one cell would be expected to be very rare. That makes unilateral retinoblastoma the expected outcome; the age of onset is generally later than in cases with inherited *RB1* mutations.

Note that while people with bilateral tumors can be confidently expected to have inherited a germline *RB1* mutation, a minority of people with unilateral tumors also have a germline mutation (by chance, tumor formation has only occurred in one eye). Inheritance of retinoblastoma susceptibility is dominant, but incompletely penetrant.

The two-hit paradigm explains why the cancer can be transmitted in a dominant fashion, even although the phenotype is recessive at the cellular level (both alleles of a tumor suppressor gene need to be inactivated or silenced); see Figure 10.10B. It also applies well to some other cancers that exist in both familial and sporadic forms but not to some other cancers, as described below.

Loss of heterozygosity

For tumor suppressor genes, the initial hit is typically confined to the tumor suppressor locus—usually an inactivating point mutation. Inactivation of the second allele can also occur by a locus-specific DNA change—a point mutation, gene deletion, or gene conversion—or sometimes by epigenetic silencing.

Often, however, the second allele is inactivated by large-scale DNA changes (loss of the whole chromosome, or a substantial part of it) or by mitotic recombination (Figure 10.11). In that case *loss of heterozygosity* will be evident: linked DNA markers that are constitutionally heterozygous in normal blood cells from an individual contain just a single allele in the tumor sample.

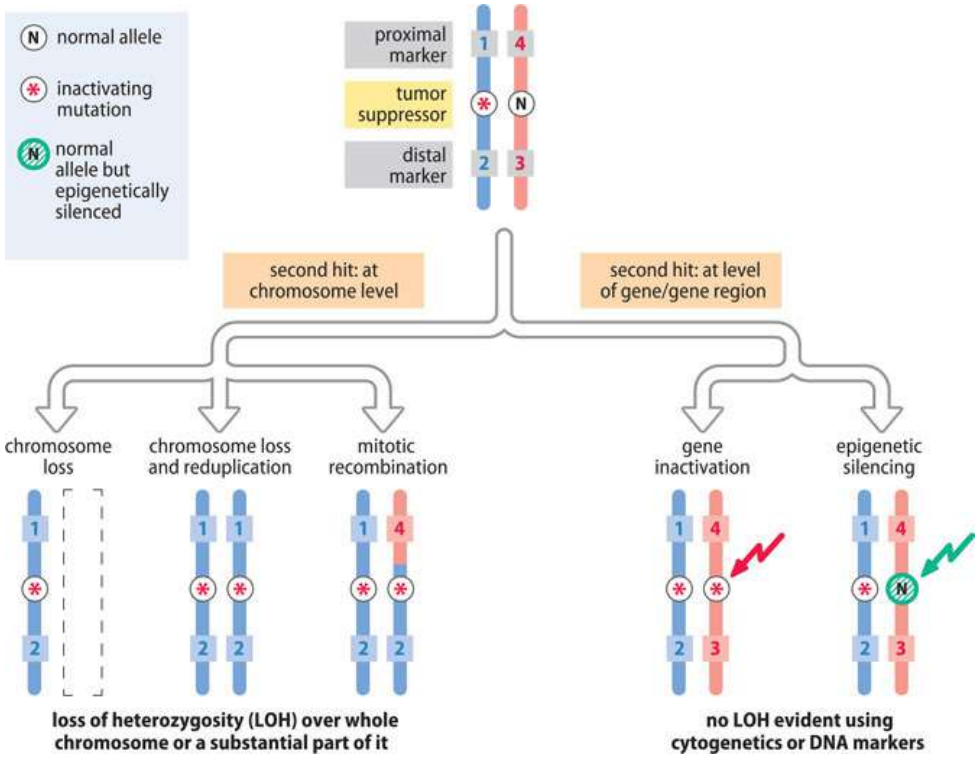


Figure 10.11 Different types of second hit at a tumor suppressor locus, some readily detectable by screening for loss of heterozygosity, and others not. Here the first hit at the tumor suppressor locus is shown as a small-scale inactivating mutation on the blue chromosome. A second hit that involves a large-scale (chromosomal) change (such as loss of the orange chromosome or loss of a part of that chromosome by mitotic recombination) can result in obvious *loss of heterozygosity* (readily detectable at the level of cytogenetic or DNA marker analysis). Sometimes, however, the second hit can be an inactivating mutation at the second allele or an epigenetic silencing event encompassing the tumor suppressor locus. In these cases,

both alleles are unable to be expressed but loss of heterozygosity would not be evident by either cytogenetic analyses or DNA analyses using flanking markers.

Loss of heterozygosity has been used as a way of mapping tumor suppressor genes. Paired samples of blood and the relevant tumor from individuals are screened with DNA markers from across the genome to identify chromosomes and, more profitably, chromosomal regions that show convincing loss of constitutional marker heterozygosity in the tumor samples. Analysis of multiple different tumors might lead to the identification of a quite small subchromosomal region defined by different mitotic crossovers or other breakpoints observed.

The key roles of gatekeeper tumor suppressor genes in suppressing G_1-S transition in the cell cycle

Understanding how cell division is regulated is of paramount importance in understanding cancer. Protein complexes made up of cyclins and cyclin-dependent kinases (CDKs) have key roles in regulating the cell cycle at certain cell cycle *checkpoints*.

The regulation of G_1, the phase of the cell cycle when cells make the decision whether or not to divide, is pivotal in tumorigenesis. A principal checkpoint occurs late in G_1, close to the G_1/S boundary, and is subject to intense regulation. A complex of cyclin E and the CDK2 protein works at this checkpoint to promote the transition from G_1 to S phase (which will usually commit the cell to cell division).

The CDK2–cyclin E complex is in turn regulated by interconnecting pathways. The control system has two arms in which the RB1 and p53 tumor suppressor proteins have commanding roles; another three tumor suppressor proteins, p14, p16, and p21 (the numbers refer to initially estimated molecular weights in kDa) support p53 and RB1 in putting a brake on cell division (**Figure 10.12**).

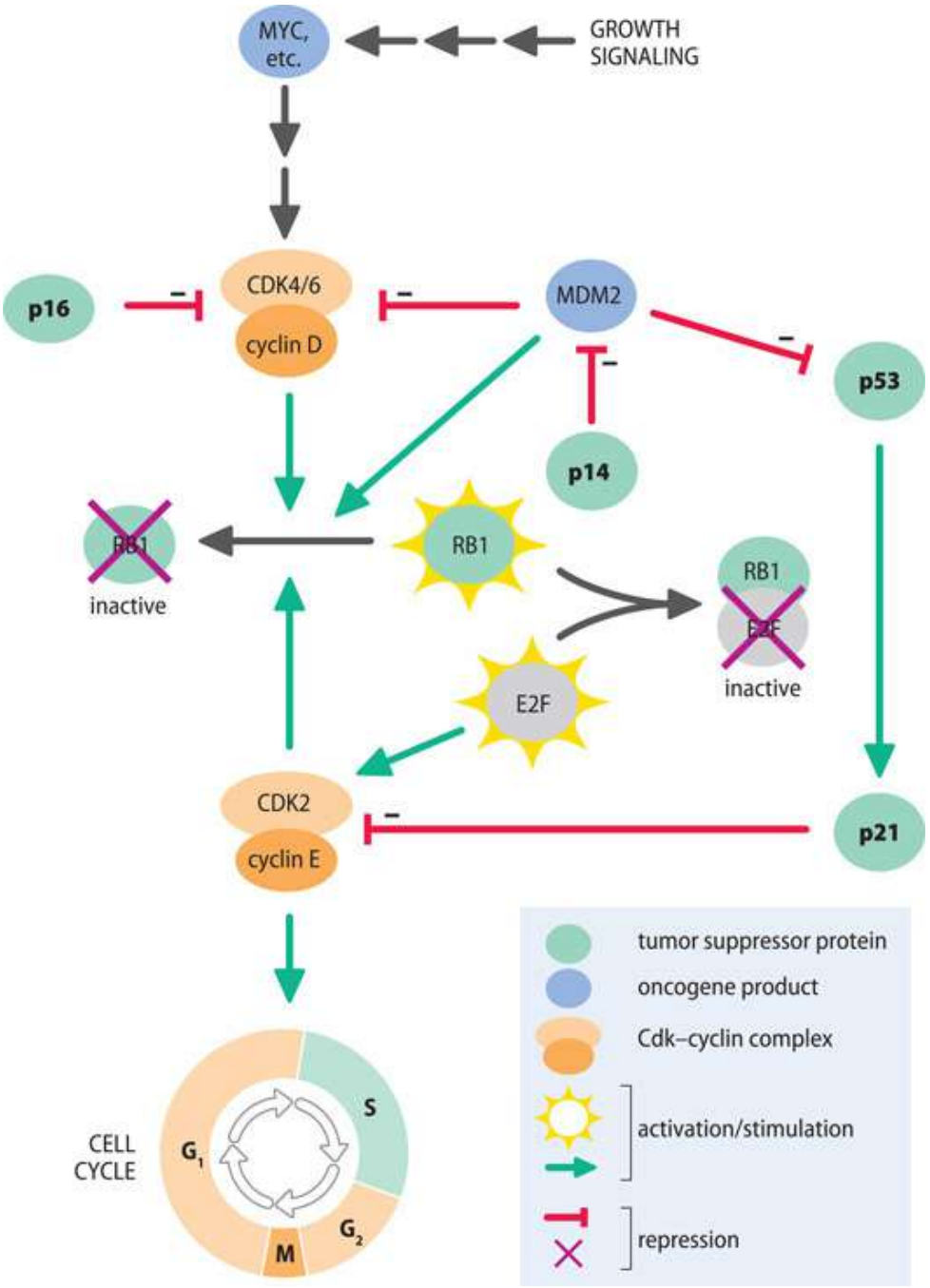


Figure 10.12 Major roles for p53, RBI, and accessory tumor suppressor proteins as brakes on cell growth. To permit cell growth, the CDK2–cyclin E complex promotes the G1-S transition and is stimulated to do so by the E2F transcription factor. (E2F activates the transcription of multiple genes whose products are required for progression to S phase, notably cyclin E.) Five tumor suppressor proteins work in the opposite direction, as brakes on cell growth. RB1 inhibits the E2F transcription factor by binding to it to keep it in an inactive form. It, in turn, is repressed by CDK4/6-cyclin D and MDM2 but is assisted by proteins that repress

its inhibitors: p16 (also called INK4a because it inhibits cyclin-dependent kinase 4) and p14 (also called ARF; it inhibits MDM2). p53 and p21 work in a pathway that bypasses E2F to inhibit the CDK2-cyclin E complex directly. Normally, p53 concentrations are kept low in cells but are increased in response to severe DNA damage. Elevated p53 both suppresses cell division (as shown here) and also stimulates apoptosis pathways (as shown in Figure 10.13).

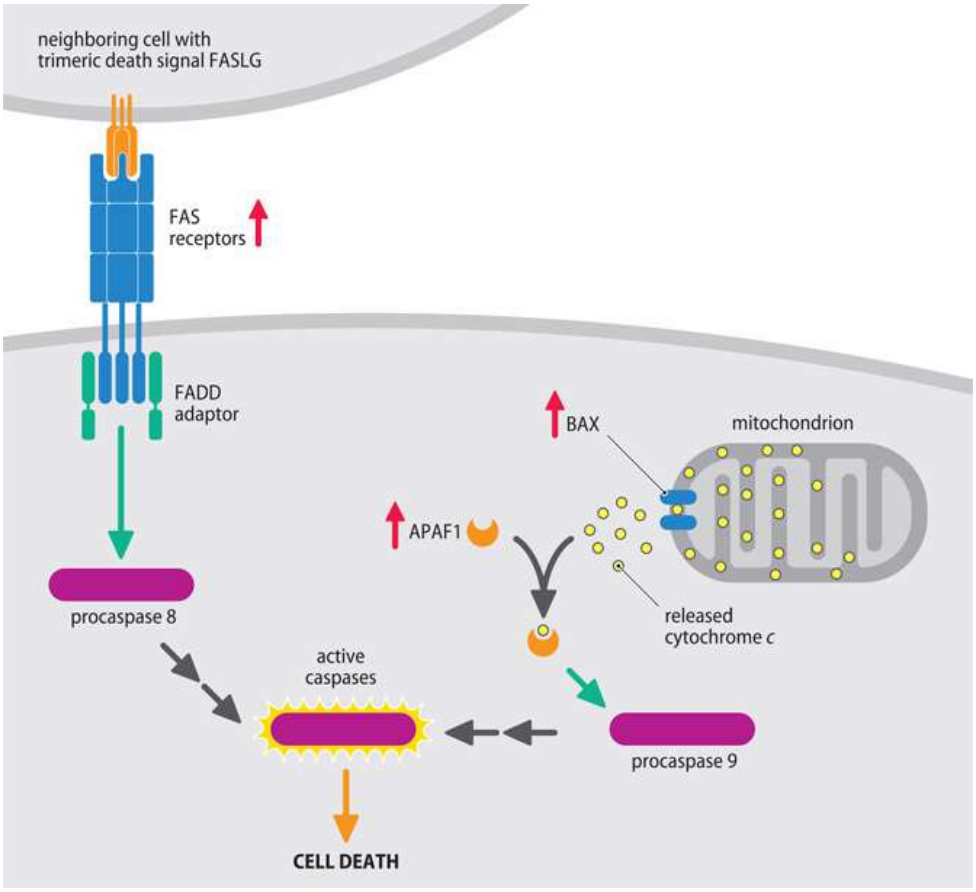


Figure 10.13 Regulation of different apoptosis pathways by p53. When actively expressed at high concentrations, p53 stimulates the transcription of various genes to produce increased quantities of apoptosis-promoting proteins (indicated by vertical red arrows). They include cell surface receptors that are able to recognize death signals from neighboring cells, such as FAS receptors, and regulators of the mitochondrial apoptosis pathway, notably BAX and APAF1. FAS receptors are monomers, but when contact is made with the trimeric FAS ligand (FASLG) they form trimers. The FAS trimers recruit an adaptor (FADD), forming a platform for binding and activating procaspase 8. The BAX1 protein forms oligomers within the mitochondrial outer membrane that act as pores, allowing the release of cytochrome *c* into the cytosol. The released cytochrome *c* binds and activates APAF1, which in turn binds and activates procaspase 9. Activated procaspases 8 and 9 ultimately lead to mature effector caspases that destroy the cell.

Growth signaling pathways can induce a loss of RB1 function by stimulating CDK4–cyclin D or CDK6–cyclin D complexes that inactivate RB1 (by phosphorylating it) and the negative regulator MDM2 (which adds ubiquitin residues to target RB1 for destruction to keep RB1 levels low when growth is needed). Otherwise, p16 and p14 work to suppress the inhibition of RB1 and so put a brake on cell growth (Figure 10.12).

Elevated levels of p53 protein can stimulate p21 to inhibit the CDK2-cyclin E complex so as to induce cell cycle arrest. However, p53 is normally kept at relatively low concentrations in cells, mostly because MDM2 binds to p53 and adds ubiquitin groups to it, targeting it for destruction.

Signals from different sensors that detect cell stress—such as sensors of DNA damage—result in phosphorylation of p53. Phosphorylated p53 is not bound by MDM2 and so p53 levels increase. This can lead to cell cycle arrest (see Figure 10.12), which provides the opportunity to repair DNA, or to apoptosis when the DNA damage is too severe to repair. As detailed below, p53 has a pivotal role in cancer and is a rather unconventional tumor suppressor.

Note that although very different in sequence, the p14 and p16 tumor suppressors are both made from alternative splicing of a single gene, *CDKN2A* (see Figure 6.8B on page 146), and loss-of-function mutations in this one gene can inactivate both the RB1 and p53 arms of the cell cycle control system. Not surprisingly, *CDKN2A* mutations are important in tumorigenesis, and homozygous deletion or inactivation of this gene is quite common in cancers.

The additional role of p53 in activating different apoptosis pathways to ensure that rogue cells are destroyed

Cells that are unwanted, heavily damaged, or actively dangerous are normally induced to commit suicide through **apoptosis** (programmed cell death) pathways. Some apoptosis pathways work through a cell surface receptor that receives a "death signal" from neighboring cells (examples include FAS receptors and other members of the tumor necrosis factor receptor superfamily). Other pathways, such as the mitochondrial apoptosis pathway, respond to certain types of internal damage, such as that caused by harmful reactive oxygen species or exposure to dangerous levels of ionizing radiation.

In most cases the apoptosis pathway ends by triggering the cell to produce certain caspases, proteolytic enzymes that wreak havoc by inactivating all kinds of

important proteins in the cell; an endonuclease is also activated that cleaves DNA into small fragments. Because each of our normal cells has the potential to commit suicide, apoptotic pathways need to be very tightly regulated.

Various cancer-associated genes make products that regulate apoptosis. They include some tumor suppressors, notably *TP53*. When an unexpected double-strand break occurs in DNA, the DNA damage response activates high-level expression of p53. In response, p53 may activate transcription of various apoptosis-promoting genes in different apoptosis pathways (**Figure 10.13**).

From the above, we can see that p53 has dual central roles: it inhibits excessive cell proliferation, and it also acts as a "guardian of the genome" by inducing apoptosis in response to double-strand DNA breaks (which are common in cancer cells). To promote cell proliferation and inhibit apoptosis, cancers frequently seek to inactivate both *TP53* alleles, and *TP53* is the most commonly mutated gene in cancer.

Note that oncogenes also have a role in inhibiting apoptosis. For example, the oncogene *BCL2* works in the mitochondrial apoptosis pathway, where its protein product inhibits cytochrome *c* release from mitochondria and is inhibited in turn by the BAX protein. In cancer cells, over-expression of certain oncogenes, such as *BCL2*, inhibits apoptosis.

Tumor suppressor involvement in rare familial cancers and non-classical tumor suppressors

Familial cancer is comparatively infrequent. Some rare examples are known of heritable oncogene mutations that cause cancer. For example, germline mis-sense mutations in the *RET* proto-oncogene are found in familial thyroid cancer. However, the great majority of familial cancers have germline mutations in tumor suppressor genes, including both gatekeeper genes such as *RB1* (with normal roles in restraining cell proliferation and/or promoting apoptosis) and caretaker genes (with genome maintenance roles, notably in DNA repair); **Table 10.5** gives some examples.

TABLE 10.5 EXAMPLES OF FAMILIAL CANCERS RESULTING FROM GERMLINE MUTATIONS IN TUMOR SUPPRESSOR GENES

Familial cancertype	Gene*	Normal function of gene product(s)

Familial cancertype	Gene*	Normal function of gene product(s)
DEFECT IN GATEKEEPER GENE		
Familial adenomatous polyposis coli	*APC*	multiple functions, notably in signal transduction (Wnt pathway)
Familial melanoma	*CDKN2A*	two unrelated protein products, p14 and p16, facilitate p53-mediated cell cycle arrest (see Figure 10.12)
Gastric carcinoma	*CDH1*	regulator of cell-cell adhesion
Gorlin syndrome (basal cell carcinoma, medulloblastoma)	*PTCH*	sonic hedgehog receptor
Juvenile polyposis coli	*DPC4*	signal transduction (TGFβ pathway)
	SMAD4	
Li-Fraumeni syndrome (multiple different tumors)	*TP53*	the p53 transcription factor induces cells to undergo cell cycle arrest (see Figure 10.12) or apoptosis (see Figure 10.13)
Neurofibromatosistype 1 (NF1)	*NF1*	negative regulation of Ras oncogene
Neurofibromatosis type 2 (NF2)	*NF2*	cytoskeletal protein regulation
Retinoblastoma	*RB1*	acts as a brake on the cell cycle (see Figure 10.12)
Wilms tumor (childhood kidney tumor)	*WT1*	a transcriptional repressor protein with multiple functions including regulating the fetal mitogen insulin-like growth factor
DEFECT IN CARETAKER GENE		
Familial breast/ovarian cancer	*BRCA1*	makes product that interacts with double-strand DNA repair complex/components
	BRCA2	
Hereditary non-	*MLH1*	DNA mismatch repair

Gatekeeper genes include classical tumor suppressors that work in regulating cell division or upstream growth signaling pathways. Caretaker genes include other tumor suppressors that work in DNA repair or DNA damage responses.

* Predisposing locus that shows germline mutations.

Familial cancertype	Gene*	Normal function of gene product(s)
polyposis colorectal cancer (Lynch syndrome)	*MSH2*	

Gatekeeper genes include classical tumor suppressors that work in regulating cell division or upstream growth signaling pathways. Caretaker genes include other tumor suppressors that work in DNA repair or DNA damage responses.

* Predisposing locus that shows germline mutations.

In retinoblastoma, few driver mutations are thought to be required for tumorigenesis (embryonic retinal progenitor cells are both poorly differentiated and rapidly proliferating, and so these cells already have two important tumor cell characteristics). The two-hit paradigm applies to additional types of cancer in which investigation of rare familial forms led to the identification of a tumor suppressor gene that was then found to be mutated in more common sporadic forms.

Some cancers that exist in both heritable and sporadic forms do not, however, readily fit the classical two-hit tumor suppressor paradigm. Major tumor suppressor genes implicated in the common sporadic tumors are often different from those involved in familial forms. This can be explained at least in part by disease heterogeneity. For example, *BRCA1,* the principal tumor suppressor gene implicated in familial breast cancer, is inactivated in only 10–15 % of sporadic breast cancers. The latter form a recognizably distinct subset of sporadic breast cancers (and in these cases any second hit occurs by epigenetic silencing). Other data from many cancers have prompted the need for a radical overhaul of the classical two-hit suppressor hypothesis, as described in the next subsections.

Non-classical tumor suppressors

Some cancer-susceptibility genes seem to lose the function of one allele but the second allele seems perfectly normal at the DNA level. Sometimes the second allele is *epigenetically silenced*. In other cases, however, inactivating a single allele seems to be sufficient to induce a tumorigenic change; that is, a significant contribution to tumorigenesis can be made by heterozygous loss of function, or *haploinsufficiency*.

Examples include some tumor suppressor genes involved in genome stability for which homozygous inactivation would be expected to lead to cell death (but can be averted by a third hit such as mutation in *TP53).* And mutation of just a single allele of certain tumor suppressor genes, such as *BRCA1*, has been shown to lead to genome instability in cultured cells and animal models.

Gain-of-function mutations can also occur in some tumor suppressor genes; in that case, a single mutated tumor suppressor allele can behave like an onco-gene. For example, missense mutations are very common in *TP53,* which makes the p53 tumor suppressor, and the resulting mutant p53 proteins can behave in a dominant-negative fashion (**Box 10.3**).

Partial inactivation of tumor suppressors can also make vital contributions to tumorigenesis; even quite subtle changes to the dosage of some tumor suppressors can sometimes make a substantial difference. The dosage effects can be highly tissue-specific and dependent on the context, such as the genetic background—for the example of the *PTEN* tumor suppressor, see the review by Berger et al. (2011) under Further Reading.

BOX 10.3 A CENTRAL ROLE IN CANCER FOR THE *TP53* SUPPRESSOR GENE THAT MAKES A NON-CLASSICAL TUMOR SUPPRESSOR, p53

The *TP53* gene at 17p13 has a central role in cancer, being mutated in nearly half of all tumors. The gene product, p53, has many roles and is involved in numerous different features of cancer. However, much of its importance comes from its role as a "guardian of the genome"—it connects DNA damage, a common feature in cancer cells (which frequently undergo genome instability as described in Section 10.3), to decisions to induce cell cycle arrest (see Figure 10.12) or apoptosis (see Figure 10.13).

The p53 control mechanism that seeks to nip tumor-igenesis in the bud can never be a failsafe mechanism; as a back-up, two p53-related proteins, p63 and p73, are produced with functions that partly overlap those of p53. Nevertheless, p53 has the dominant role.

As befits its crucial role, p53 is expressed in all cells. Germline mutations in *TP53* underlie Fraumeni syndrome (OMIM 151623), a dominantly inherited

disorder in which those affected within a family can present with different early-onset tumors (Figure 1).

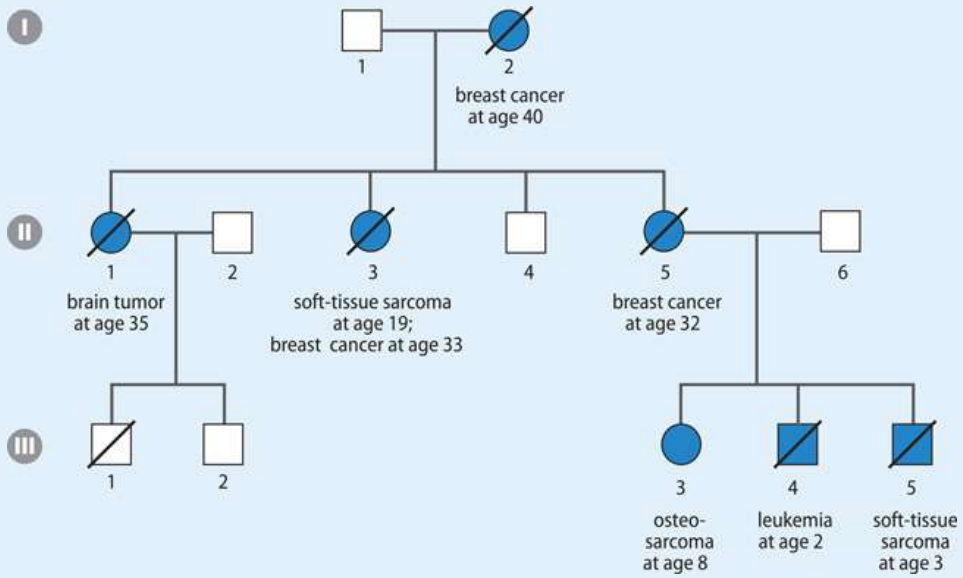


Figure 1 A typical pedigree of Li–Fraumeni syndrome. (Adapted from Malkin D [1994] *Annu Rev Genet* 28:443−465; PMID 7893135. With permission from Annual Reviews.)

p53 AS A NON-CLASSICAL TUMOR SUPPRESSOR

In several ways, p53 does not behave as tumor suppressor. In most tumor suppressors (such as *RB1, APC, NF1, NF2,* and *VHL)* the primary mutations are mostly deletion or nonsense that result in little or no expression of the respective proteins. *TP53* is different: the great majority of small-scale cancer-associated mutations are single nucleotide missense mutations that are very largely clustered within the central DNA-binding domain.

Six codons are predominantly mutated within the DNA-binding domain, and the missense mutations fall into two classes (Figure 2). In the DNA contact class, the missense mutation alters an amino acid that is normally used to make direct contact with the DNA of genes regulated by p53. The conformation class of mutations disrupt the structure of the p53 protein.

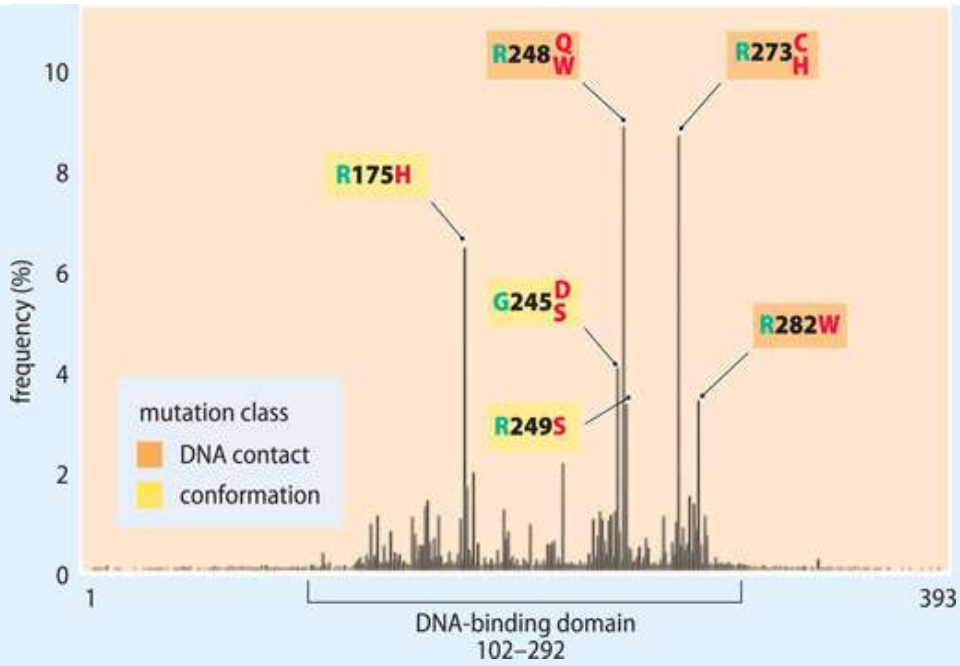


Figure 2 ***TP53*** **missense mutations are very largely confined to the DNA-binding domain, with six hotspots.** Verticle black lines indicate frequencies of missense mutations at each of the 393 codon positions. Two types of amino acid replacement are seen at codons 245, 248, and 273 (for example, at codon 273 arginine is replaced by cysteine or histidine). (Adapted from Freed-Pastor WA & Prives C [2012] *Genes Dev* 26:1268–1286; PMID 22713868. With permission from Cold Spring Harbor Laboratory Press.)

The mutated p53 proteins have multiple properties that distinguish them from wild-type p53. First, unlike wild-type p53, mutant p53s do not participate in a self-limiting regulation. In normal cells, the amount of p53 is kept low because p53 is negatively regulated by MDM2 (and MDM4), and p53 positively regulates the production of its major antagonist MDM2; in cells with missense mutations in *TP53*, large amounts of mutant p53 are produced because mutant p53 fails to stimulate the production of MDM2. Mutant p53 can work in a dominant-negative fashion. It suppresses wild-type p53 and also the related p63 and p73 transcription factors (which show high sequence homology to p53 in some domains), and it antagonizes the interaction of wild-type p53 and the recognition sequences it must bind in its target genes (**Figure 3**). Instead, mutant p53 works as a rather different type of transcription factor by stimulating transcription of quite different target genes, including many genes that stimulate cellular proliferation or that inhibit apoptosis; see the review by Freed-Pastor & Prives (2012) under Further Reading.

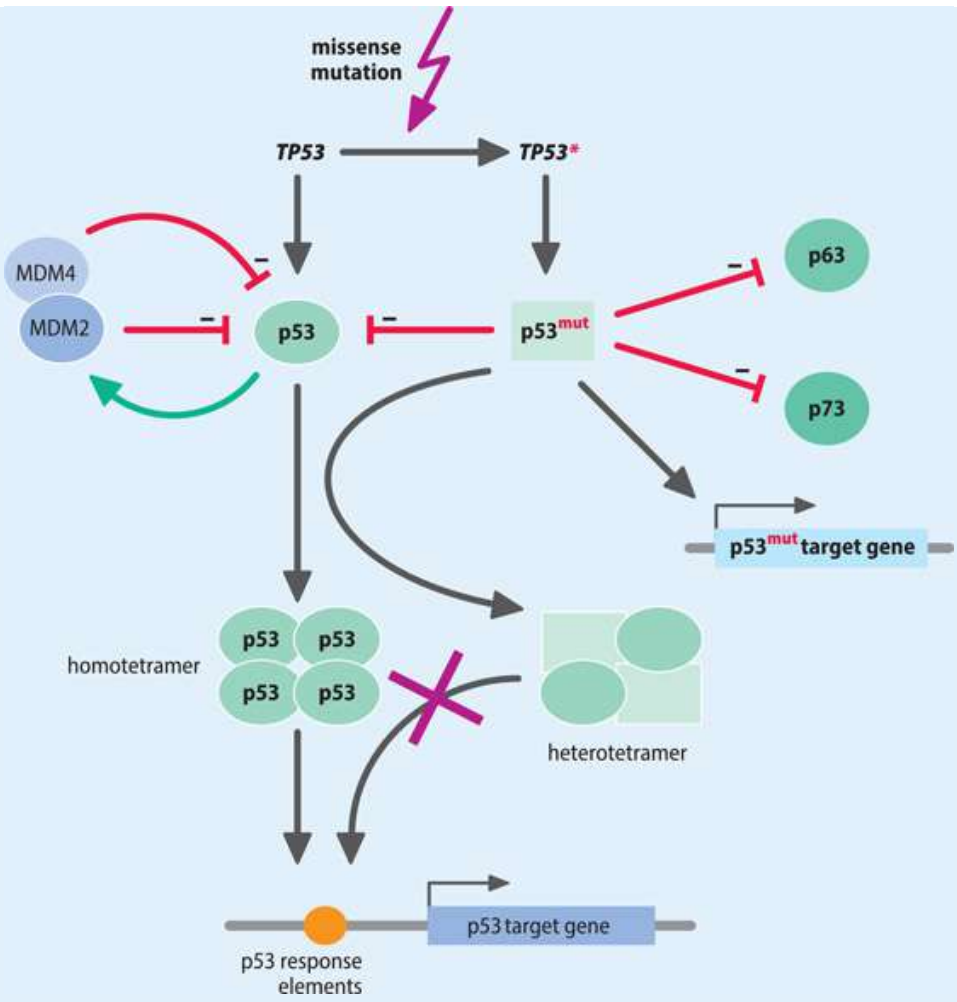


Figure 3 Missense p53 mutants have multiple novel properties and can show dominant-negative interactions with wild-type p53. Wild-type p53 works as a homotetramer to recognize and bind DNA sequences with specific motifs (p53 response elements) in the control regions of the p53 target genes. The p53 missense mutants suppress both wild-type 53 and the related p63 and p73 transcription factors. Mutant p53 is produced in very large amounts (unlike wild-type p53, it is not subject to self-regulation through stimulation of the MDM2 repressor) and interferes with normal p53-mediated transcription by interacting with wild-type 53 to form unproductive heterotetramers. Instead, mutant p53 stimulates the transcription of different genes.

The significance of miRNAs and long noncoding RNAs in cancer

Irrespective of the class of cancer-susceptibility gene, the normal products are almost always proteins. However, hundreds of different noncoding RNAs are also

known to be aberrantly expressed in cancer, and not surprisingly, given the widespread involvement of miRNAs in controlling gene expression, aberrant miRNA expression is very common in cancer.

Certain miRNAs are known to have important regulatory roles in processes relating to cancer, such as cell cycle control, cellular senescence, apoptosis, and DNA damage responses. Dysregulation of miRNA expression is frequent in cancer; certain miRNA genes can be lost in cancer cells and other miRNAs are known to be overexpressed in certain tumors.

On the basis of the above observations alone, various miRNAs have been viewed as behaving as tumor suppressors or oncogenes. For example, the *MIR15A* and *MIR16–1* genes at 13q14 have been regarded as tumor suppressors on the basis that they normally induce apoptosis by targeting BCL-2, but are frequently deleted or down-regulated in chronic lymphocytic leukemia. The *MIR21* gene, which regulates the *PTEN* and *PDCD4* genes and is over-expressed in many solid tumors, might be an example of a miRNA gene that can behave as an oncogene.

The dysregulated expression of miRNA and the loss of certain miRNA genes in cancers may be important events in cancer progression. However, it is difficult to evaluate the contributions of individual miRNAs: a single miRNA can regulate many different mRNAs, and a single mRNA may be regulated by many different miRNAs. At the time of writing, there seems little direct evidence that miRNA genes have unambiguous roles as oncogenes or tumor suppressor genes. Nevertheless, miRNA expression patterns may help us to dissect different disease subgroups, and there is interest in using miRNAs as therapeutic targets and as cancer biomarkers.

More recently, the possible roles of long noncoding RNAs in cancer have begun to be investigated. They have been less well studied (and, for example, have not been covered in whole exome sequencing studies), but there is strong evidence that some are important in cancer. In mice, for example, the *Xist* gene is not just involved in X-chromosome inactivation but also suppresses cancer *in vivo;* if *Xist* is deleted in blood cells, mutant females develop a highly aggressive myeloproliferative neoplasm and myelodysplastic syndrome with 100 % penetrance. Readers who may be interested in the role of noncoding RNAs in cancer can find a recent review at PMID 31730848.

10.3 GENOMIC INSTABILITY AND EPIGENETIC DYSREGULATION IN CANCER

The evolution of cancer cells is driven not just by a series of changes to the genome, but also by epigenetic changes. Natural selection at the cell level drives cells to relax normal controls on cell proliferation and apoptosis. Achieving this involves efforts to subvert both genome and epigenome stability. As we describe below, epigenetic changes may sometimes initiate the process of tumorigenesis.

An overview of genome and epigenome instability in cancer

Genome instability is an almost universal characteristic of cancer cells, and frequently results from defects in chromosome segregation or DNA repair. By weakening the capacity to maintain the integrity of the genome, more DNA changes will be generated for natural selection to work on to drive tumor formation. Eventually, a cell can build up a sufficient number of DNA changes to become an invasive cancer cell. Genomic instability can manifest itself at two levels:

- *at the chromosomal level*. Chromosomal instability (sometimes abbreviated as CIN) is a particularly common form of genome instability. Tumor cells typically have grossly abnormal karyotypes (extra or missing chromosomes and many structural rearrangements), and they often show chromosomal instability in culture
- *at the DNA level.* The instability may be genome-wide or localized. As detailed below, a genome-wide form of DNA instability is especially evident in some types of colon cancer. Sporadic colorectal tumors either show chromosome instability (in most cases) or global DNA instability (in about 15 % of cases), but not both: the instability seems to be the result of natural selection. More localized DNA instability is exemplified by the phenomenon known as *kataegis, a* form of clustered hypermutation first reported in 2012. We provide details within the context of cancer genomics in Section 10.4.

Epigenetic dysregulation is a feature of all cancer cells, ranging from apparently normal precursor tissue to advanced metastatic disease. As well as being important in cancer progression (and helping cancer cells achieve each of the 10 characteristic biological capabilities listed in Table 10.2 above), epigenetic dysregulation can be a key step in the initiation of cancer. And as we describe below, certain types of epigenetic dysregulation also cause chromosome instability and accelerated genetic changes in tumor cells.

The reader may justifiably wonder how cancer cells with their often bizarre chromosome constitutions, plus both DNA and epigenetic stability and inefficient anaerobic glycolysis, would be able to grow and proliferate as much as they do. But somehow these features are the outcome of the relentless drive by natural selection to enable cells to escape many layers of controls on cell division and programmed cell death.

Different types of chromosomal instability in cancer

Chromosome abnormalities are important in accelerating tumorigenesis: oncogenes can be activated by rearrangements such as translocations, and tumor suppressor alleles can be lost through deletions, whole chromosome loss, or recombinations. Standard cytogenetic methods are often difficult to carry out on tumor cells, but various DNA-based methods can be used to study chromosome instability in cancer, and they can have quite a high resolution. They include two microarray-based DNA hybridization methods—one based on comparative genome hybridization, and one on SNP (single nucleotide polymorphism) analyses—which we introduce within the general context of DNA-based diagnosis in Section 11.1. Another method is *spectral karyotyping,* a type of multicolor chromosome FISH (fluorescence *in situ* hybridization). Unlike the microarray hybridization-based methods, it can reveal balanced chromosome abnormalities (in which there is no net loss or gain of DNA), as well as unbalanced chromosome abnormalities; see **Figure 10.14** for an application.

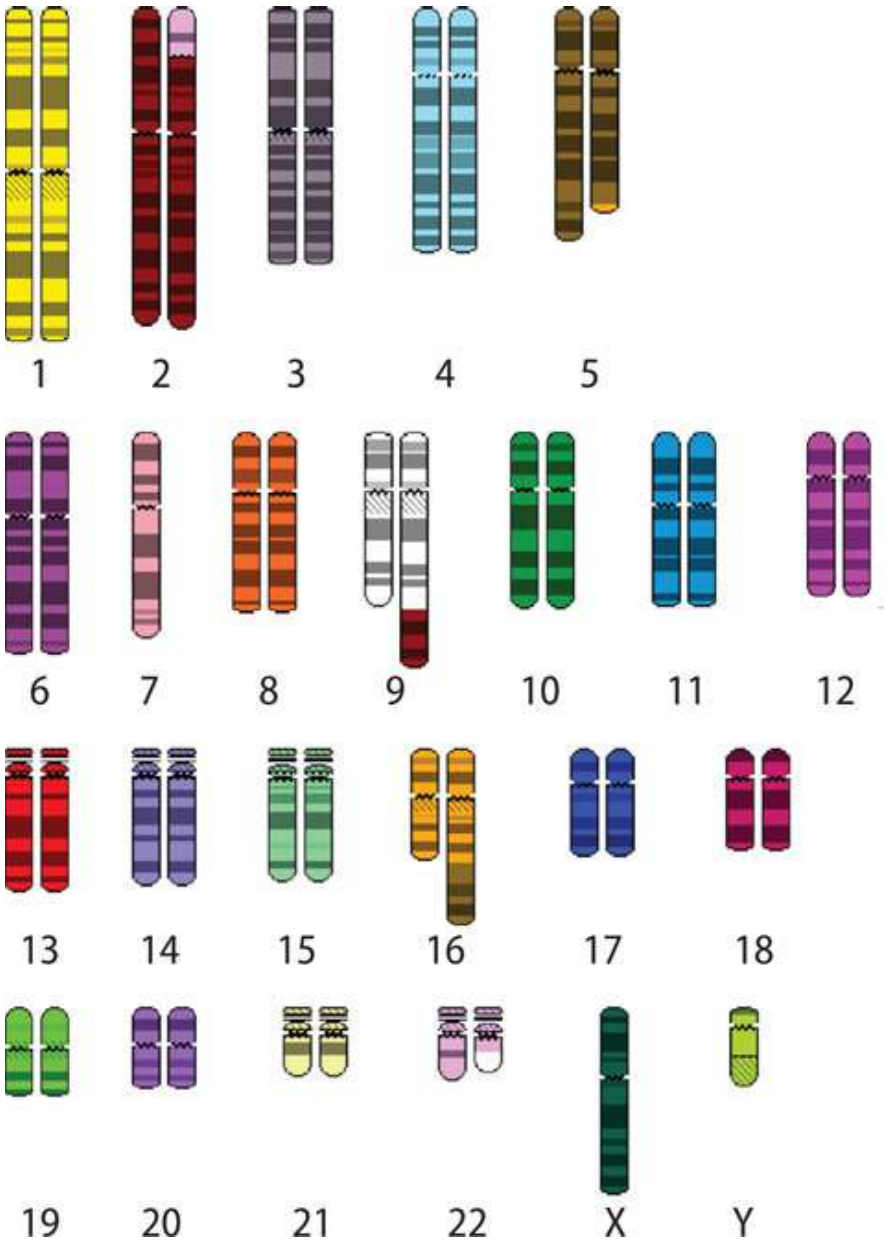


Figure 10.14 An example of using spectral karyotyping to analyze chromosomes in tumor cells. Spectral karyotyping (SKY) is a variant of chromosome fluorescence *in situ* hybridization (FISH) in which cocktails of many fluorescently labeled DNA probes from different regions of chromosomes are used to "paint" chromosomes so that entire chromosomes are labeled with a specific fluorochrome and become fluorescent. Different chromosomes are painted with different combinations of multiple fluorescent labels. An image analyzer scans the fluorescent signals and can discriminate between the different fluorescence signals used for each of the 24 different chromosomes. To help us visualize the result it assigns artificial ("false") colors for each chromosome signal. In this example, there is a three-way variant of the standard 9;22 translocation (involving chromosome 2), plus an additional 5;16 translocation and the loss of one

copy of chromosome 7. The karyotype is interpreted as 45,XY,t(2;9;22) (p21;q34;q11),t(5;16) (q31;q24),–7.

A major source of aneuploidies in cancer cells is defects in the spindle checkpoint, the cell cycle control mechanism that checks for correct chromosome segregation (it normally ensures that the anaphase stage of mitosis cannot proceed until all chromosomes are properly attached to the spindle). Extra centrosomes are often seen in cancer cells and may trigger the formation of abnormal spindles and unequal segregation of chromosomes into the daughter cells.

Structural chromosome abnormalities in cancer cells can arise in different ways. The most common source is an abnormal response to unrepaired DNA damage. As detailed in Section 4.2 we have complex DNA repair systems that can never be 100 % efficient. Normally DNA damage responses act as a backup: they trigger apoptosis if the DNA damage is severe, or they arrest the cell cycle so that an unrepaired defect can be repaired. Defects in DNA repair of DNA damage responses allow unrepaired or damaged DNA to be passed on to daughter cells.

Failure to repair double-strand DNA breaks is an important source of structural chromosome abnormalities and can be precipitated by inactivation of key caretaker genes that function in this repair pathway, including the breast-cancer-associated *BRCA1* and *BRCA2* genes, which work in homologous recombination-mediated DNA repair. Proteins such as the ATM (ataxia–telangiectasia mutated) protein kinase work as sensors to detect unprogrammed double-strand DNA breaks. They then activate signaling mediators, which in turn recruit effectors to repair the damage. As well as DNA repair, the DNA damage response involves arresting the cell cycle, notably by activating p53 (**Figure 10.15**).

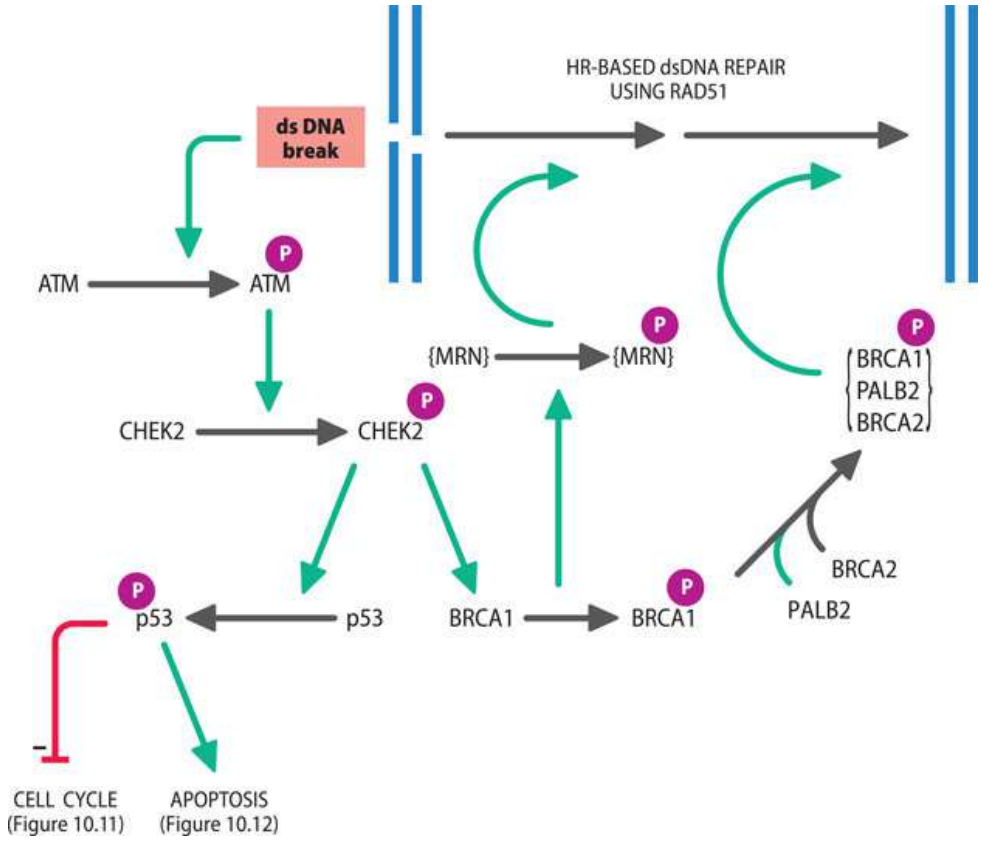


Figure 10.15 Some cellular signaling responses to double-strand DNA breaks and different roles of BRCA1 and BRCA2 in homologous recombination-based DNA repair. Homologous recombination (HR) appears to be the major mechanism for repairing double-strand breaks in proliferating cells. Green arrows indicate stimulatory reactions; the red T-bar indicates inhibition. The ATM protein kinase is a prominent sensor of DNA damage. It is activated by phosphorylation (P), and in turn causes phosphorylation-mediated activation of CHEK2, which similarly activates p53 and BRCA1. Phosphorylated BRCA1 has multiple roles including activating protein complexes that are directly involved in HR-mediated double-strand (ds) DNA repair. These complexes—shown here by curly brackets—include the MRE11–RAD50–NIBRIN (MRN) complex, and also a complex in which activated BRCA1 recruits BRCA2 through an intermediary binding protein PALB2. As well as activating DNA repair, the DNA damage response may initiate cell cycle arrest, notably by activating p53 (which works at the G_1–S checkpoint); if DNA damage cannot be readily repaired, it can also activate apoptosis through enhanced p53 production.

Chromothripsis and chromoplexy

Sometimes chromosome breakage can involve an extensive localized rearrangement of chromosomes. In the process of *chromothripsis* large numbers of chromosomal rearrangements are generated in what appear to be single catastrophic events. The chromosome rearrangements may occur by chromosome shattering and aberrant rejoining of fragments by error-prone end- joining DNA repair pathways, or by aberrant DNA replication-based mechanisms. Chromothripsis may not be common in many cancers, but it is significantly more frequent in cells with mutated p53. Interested readers can find a recent review at PMID 28899600.

Another, somewhat bizarre type of chromosome rearrangement, known as *chromoplexy*, can occur in tumors where chains of linked chromosomes form by serial translocations. An initial translocation might arise between chromosomes A and B, but unlike in conventional reciprocal translocation, the broken ends are not joined together to form hybrid A–B and B–A chromosomes. Instead, they may engage in new translocation with chromosomes C and D, which then generates another pair of broken ends that engage in further translocation with chromosomes E and F and so on. Interested readers can find a recent review at PMID 23680143.

Telomeres and chromosome stability

In human cells, the telomeres shorten at cell division (usually by about 30–120 nucleotides at each cell division). By inactivating normal controls on cell growth, cancer cells can reach a stage where some telomeres become so short that the cell can misinterpret the ends of seriously shortened telomeres as breaks in double-strand DNA. That alerts a DNA repair pathway that attempts a repair by fusing chromosomes at their ends. The resulting chromosomes with two centromeres may be pulled in opposite directions at mitosis, causing further broken ends and new cycles of chromosome fusion and breakage. Cancer cells seek to avoid this type of chromosome instability, and most frequently the solution involves ensuring that telomerase somehow becomes expressed (as previously detailed in Box 10.1).

Deficiency in mismatch repair results in unrepaired replication errors and global DNA instability

Mutation in genes involved in different types of DNA repair leads to cancer. Defective homologous recombination-based DNA repair is associated with various types of cancer, notably breast, ovarian, and pancreatic cancer. Genetic deficiency in components of nucleotide excision repair produces syndromes with increased cancer susceptibility, notably xeroderma pigmentosum (OMIM 278700). Genetic defects in base excision repair are associated with certain neurological disorders and occasionally cancer.

Germline mutations in both copies of the *MUTYH* gene (which is involved in the repair of adenines that are inappropriately base paired with guanine, oxoguanine, or cytosine) result in an autosomal recessive form of familial adenomatous polyposis (FAP), a type of hereditary colon cancer in which multiple polyps (adenomas) develop. Deficiency in **mismatch repair**, which corrects errors in replication that for some reason have not been detected by the proofreading activity of a DNA polymerase, results in a global form of DNA instability and is most commonly associated with colon cancer.

The mismatch repair mechanism

The mismatch repair (MMR) components work closely with the DNA replication machinery. In human cells, three types of protein dimer carry out most of the repairs (**Figure 10.16A**). Two of them—hMutSa and hMutsb—are needed to identify base mismatches. hMutSa identifies base-base mismatches but can also handle mismatching due to single-nucleotide insertions or deletions. hMutSb can spot base mismatching for different sizes of very short insertions or deletions (which frequently occur at short tandem repeats as a result of replication slippage, the tendency for DNA polymerase to stutter or skip forward at tandem repeats as shown in Figure 4.6).

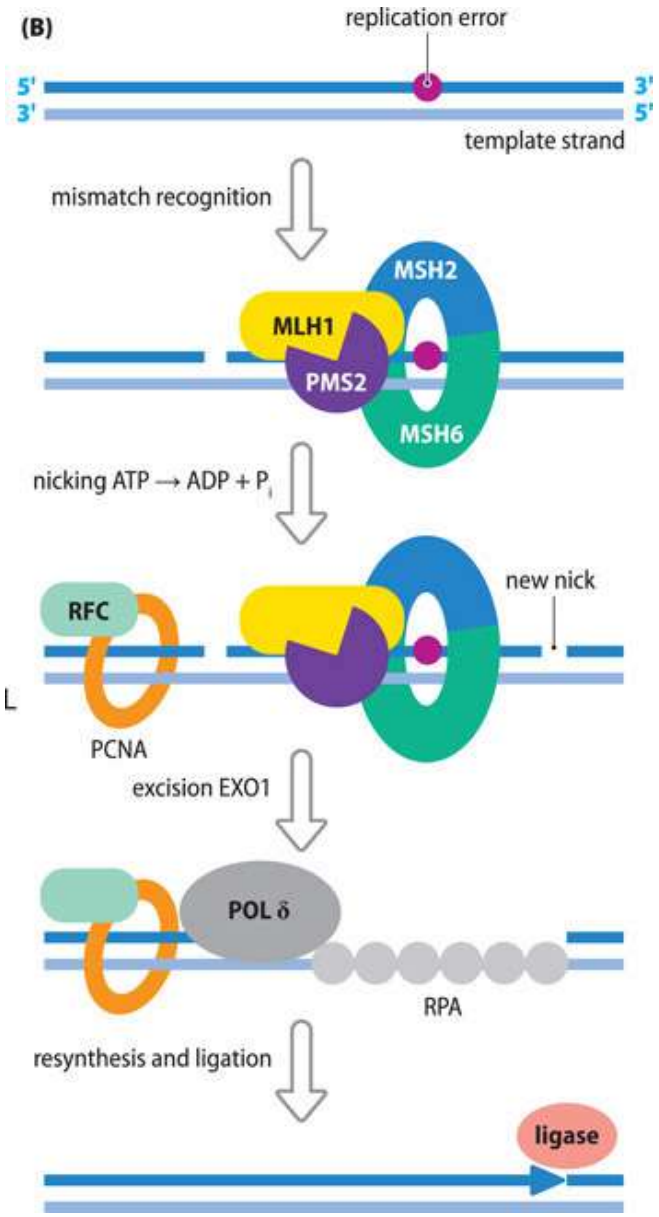


Figure 10.16 **Mismatch repair for correcting replication errors**. (A) Major classes of MutS or MutL dimers in human mismatch repair. (B) Mechanism of 5¢-directed mismatch repair in eukaryotic cells. Replication errors on a newly synthesized strand result in base mismatches that can be recognized by a MutS–MutL complex. The MutS component works as a clamp that can slide along the DNA, allowing it to scan for a base–base mismatch (MutSa) or unpaired insertion/deletion loop (often MutSb). MutLa, which has an endonuclease function, can form a ternary complex with MutS and DNA. After the newly replicated DNA has been identified (by having a preexisting nick in the DNA), PCNA (proliferating cell nuclear antigen) and RFC (replication factor C) are loaded onto the newly replicated DNA, where they help trigger the endonuclease function of PMS2 to make a new nick close to the replication error. EXO1

exonuclease is recruited to excise the sequence containing the replication error, making a gapped DNA. The resulting stretch of single-stranded DNA (stabilized by binding the RPA protein) is used as a template for the resynthesis of the correct sequence using high-fidelity DNA polymerase δ, followed by sealing with DNA ligase I. (Adapted from Geng H & Hsieh P [2013] In *DNA Alterations in Lynch Syndrome: Advances in Molecular Diagnosis and Genetic Counseling* [M. Vogelsand, ed.]. With permission from Springer Science and Business Media.)

The MMR machinery cannot simply repair one of the two strands at random: there has to be a way of distinguishing the original (correct) strand from the newly replicated strand with the incorrect sequence that needs to be repaired. Before being repaired by DNA ligase, nicks (single-strand breaks) are common on a freshly replicated DNA strand, and in human (and eukaryotic) cells the strand distinction is achieved by identifying a nearby nick on the newly replicated DNA strand. Then hMutLa cleaves the newly replicated strand close to the mismatch and recruits an exonuclease to excise a short stretch of DNA containing the replication error so that the DNA can be resynthesized and repaired (Figure 10.16B).

Consequences of defective mismatch repair (MMR)

Loss of function for both alleles of a mismatch repair gene can result in a form of global DNA instability (in which replication errors in newly synthesized DNA go uncorrected). That may be apparent in some tumors and can readily be detected by testing for global *microsatellite instability*. To do this, a selection of standard microsatellite DNA markers from across the genome are tested in tumor DNA to see if they show higher frequencies than normal for minor additional bands. Tumors demonstrating **mi**cro**s**atellite **i**nstability are described as being MSI-positive (or sometimes MIN-positive) tumors.

Lynch syndrome (also called hereditary nonpolyposis colon cancer; OMIM 120435) is a type of familial cancer in which an inactivating allele in a mismatch repair gene is transmitted by heterozygotes, predisposing to colorectal cancers and certain other cancers, including cancers of the endometrium. When investigating possible Lynch syndrome cases, if DNA analyses have not revealed an immediately obvious mutation, back-up immunohistochemistry analyses can be used as in the case study reported in **Clinical Box 13**.

In cells in which the MMR machinery is defective, the mutation rate increases about 1000-fold and so generates large numbers of mutations to drive carcinogenesis. In coding sequences, inefficient repair of base mismatches and replication slippage errors can result in gene inactivation or mutant proteins: long runs of a single nucleotide are particularly vulnerable to frameshifting insertions or deletions, and nucleotide substitutions can result in nonsense or missense mutations.

Defective mismatch repair can occur occasionally in other types of tumor, but it is particularly associated with colon cancer. Why should that be? One explanation is that MMR deficiency sabotages a key defense system that protects against colorectal cell proliferation. In the colorectum, transforming growth factor b (TGFb) is a particularly strong inhibitor of cell proliferation, and it specifically binds to a receptor on the surface of the cells of which the TGFBR2 protein is a key component. However, the *TGFBR2* gene is readily inactivated as a result of mismatch repair deficiency because it has a long sequence of adenines that make it vulnerable to frameshifting insertions and deletions (**Figure 10.17**). Somatic mutations in *TGFBR2* are found in about 30 % of sporadic colorectal cancer but are very frequent in MSI-positive colorectal cancer.

Figure 10.17 A long homopolymeric region in TGFRB2-coding DNA is a weak spot in the defense The nucleotide sequence of codons 121–130 within exon 3 of the TGFRB2 (transforming growth factor receptor β-2) gene is shown with predicted amino acids below. The sequence contains a perfect run of 10 adenines that is vulnerable to insertion or deletion by replication slippage (Figure 4.6), especially when cells become defective at mismatch repair. Resultant frameshift mutations lead to a failure to make the 567-residue TGFBR2 protein, with adverse consequences for TGFβ signaling.

CLINICAL BOX 13 CASE STUDY: LYNCH SYNDROME

Margaret presented with endometrial adenocarcinoma at 46 years of age. She had been adopted and no wider family history was available. She developed right-sided colorectal adenocarcinoma at age 47 and died at an early age. No blood sample was kept at that time, but tumor tissue sections were stored in the laboratory.

Margaret had had two children, Julia and Simon, and almost two decades after her mum died, Julia sought clinical advice in her mid-40s, concerned about the history of cancers in her mum. Analysis of archived tumor DNA samples from Margaret showed evidence of microsatellite instability suggesting a problem with mismatch repair.

A battery of tests was carried out on Julie's germline DNA sample. DNA sequencing was performed on all exons (plus flanking sequences) in the four mismatch repair genes, *MLH1, PMS2, MSH2,* and *MSH6*, and in the 3¢-UTR of the *EPCAM* gene (which neighbors the 5¢ end of the *MSH2* gene). MLPA assays (multiplex ligation-dependent probe amplification; described in Section 11.3) checked for copy number changes in relevant exons. The sequencing results did not identify a likely pathogenic point mutation, and no exon deletions or duplications were recorded by MLPA.

Subsequently, after Julia had a total hysterectomy at 46 years of age, pathology results showed an early (incidental) endometrial adenocarcinoma. Simon had no clinical symptoms but the clear family history and their mum's colorectal adenocarcinoma and tumor micro-satellite instability raised the possibility of Lynch syndrome, despite the negative DNA sequencing and MLPA results. Immunohistochemistry studies were therefore carried out on tumor samples from Julia and confirmed Lynch syndrome (see **Figure 1**).

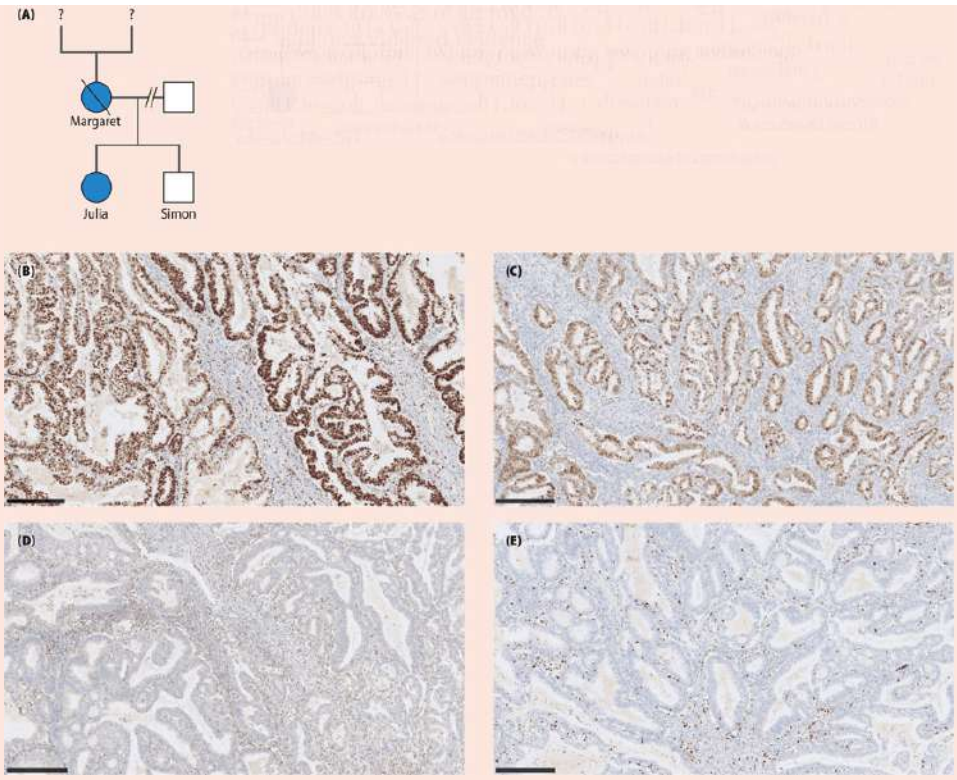


Figure 1 Family pedigree (A) and immunohistochemistry analyses on samples from Julia's tumor (B-E). Question marks indicate lack of information about Margaret's antecedents. Immunohistochemistry was performed with monoclonal antibodies specific for MLH1(B), PSM2 (C), MSH2 (D) and MSH6 (E). The strong brown staining in (B) and (C) indicates the presence of MLH1 and PMS2; the lack of brown staining in (D) and (E) shows that both MSH2 and MSH6 are not expressed.

As shown in Figure 10.16, the human MSH2 and MSH6 proteins normally work as a heterodimer (hMutSa) in mismatch repair. The two proteins are not, however, equal partners: an inactivating mutation in the *MSH2* gene that prevents production of the MSH2 protein also suppresses production of the MSH6 protein, but by contrast the MSH2 protein continues to be made after the *MSH6* gene is inactivated. Julia appears to have inherited from her mum an unidentified inactivating mutation in the *MSH2* gene, most likely a point mutation in a regulatory sequence within an intron or nearby noncoding DNA sequence. As a result, Julia was advised a clinical follow-up program with ongoing two-yearly colonoscopies.

Different classes of cancer susceptibility gene according to epigenetic function, epigenetic dysregulation, and epigenome–genome interaction

Recall that in somatic cells much of the genome is transcriptionally silenced (the heterochromatic regions and a significant, but variable, fraction of the euchromatin). This is achieved by epigenetic modifications—DNA methylation, histone modifications, and nucleosome repositioning—that attract specific proteins to compact the DNA and deny access to the transcription machinery.

The epigenetic modifications ensure that cells have distinctive chromatin patterns. They allow specific gene expression patterns to be established that determine the identity of a cell (so that it behaves as a T cell or a cardiomyocyte, for example). And they help to maintain genome stability (by maintaining the stability and function of centromeres and telomeres, and by suppressing excess activity by transposons).

One rationale for epigenetic dysregulation in tumors is that it allows cancer cells to revert to less differentiated states, permitting more flexibility to adapt to changing environments, and to assist the transformation required for the progression to cancer. It was initially thought to result simply from genetic changes in genes controlling epigenetic regulation but more recently, as detailed below, it has become clear that epigenetic changes can also initiate cancer formation.

Classifying cancer genes by epigenetic function

Previously we have classified cancer genes at two levels: by the dominant or recessive effect of genetic mutations on the phenotypes of cells (oncogenes versus tumor suppressor genes) and by selection (*driver genes*—where mutation or aberrant expression is subject to selection during tumorigenesis—and *passenger genes* that are not subject to selection towards advancing tumorigenesis). A third way of classifying cancer genes, proposed by Andrew Feinberg in 2016 (see Further Reading), is on the basis of epigenetic function, where three categories have been envisaged, as listed below.

- *Epigenetic modulator.* A gene, mutated or not, that activates or represses the epigenetic machinery in cancer. Examples include the *IDH1* and *IDH2* genes that work in the TCA (Krebs) cycle metabolism as isocitrate dehydrogenases, as detailed later. Others include the *CTCF* gene that makes a protein involved in regulating chromatin architecture and transcription, and genes making various types of cell signaling components, including the *KRAS, APC, TP53,* and *YAP1* genes.
- *Epigenetic modifier.* A gene, mutated or not, that modifies DNA methylation or chromatin structure or its interpretation in cancer. Examples include *DNMT3A* (DNA methylation), *SMARCA4* (chromatin remodeling), and *EZH2* (makes the enzymatic component of the Polycomb Repressive Complex 2 responsible for epigenetic maintenance of genes regulating development and differentiation).
- *Epigenetic mediator.* A usually unmutated gene that is regulated by an epigenetic modifier in cancer, and that increases pluripotency or survival. Examples include classic genes associated with pluripotency, such as *OCT4, NANOG, SOX2, LIN28,* and *KLF4*.

DNA methylation profiles of cancer cells and their effects on gene expression

Epigenetic profiling of cancer cells has been limited because histone modification profiles and difficult to obtain from solid tumors; for that reason, much of our information on the epigenetic profiles of tumors has come from studies on DNA methylation. In human cells, DNA methylation is almost exclusively restricted to certain cytosines that have a neighboring guanine within the dinucleotide CG (or CpG, as it is often called). Methylated cytosines can be distinguished from unmethylated cytosines by treating DNA with sodium bisulfite. (Sodium bisulfite changes all unmethylated cytosines to uracils, which become thy-mines in replicated DNA; methylated cytosines do not react and are unchanged. We describe the method in detail in Figure 11.15 on page 449.)

In somatic mammalian cells, about 70–80 % of the cytosines present in CG dinucleotides are present as 5-methylcytosine. The 5-meCG sequences are recognized and bound by specific proteins that are important in helping to organize the chromatin into compact formations that lead to transcriptional

silencing. In cancer, however, the DNA methylation patterns are changed in two ways: extensive hypomethylation and selective hypermethylation.

Across the genome as a whole, cancer cells typically show significantly reduced DNA methylation (*hypo*methylation). That includes very many genes; long blocks of sequences, enriched in repetitive DNA but containing about one-third of transcriptional start sites, are hypomethylated.

Loss of methylation in constitutively heterochromatic regions may produce aberrant transcriptional expression of highly repetitive DNA sequences, resulting in widespread chromosomal instability. That seems to be a very common event in early adenomas, for example, occurring shortly after the disturbance to the Wnt signalling pathway (mutations in *APC* or equivalent) shown in Figure 10.4B above. There is uncertainty about how this happens. One hypothesis is that the demethylation of highly repetitive DNA sequences allows normally silenced retrovirus-like elements and other related transposable elements in the genome to become active and jump to new locations in the genome, creating havoc. Constitutional DNA hypomethylation and chromosomal instability are also features of some human disorders, such as ICF1 (immunodeficiency with centromeric instability and facial anomalies; OMIM 242860), an autosomal recessive disorder that often results from mutations in the *DNMT3B* DNA methyltransferase gene.

DNA *hyper*methylation commonly occurs at the promoters of a few hundred genes in cancer cells, including tumor suppressor genes, DNA repair genes, and genes encoding certain transcription factors that are important in differentiation. Some tumor suppressor genes, such as *CDKN2A* and *MGMT*, are frequently silenced in a wide range of tumors; for others, silencing is limited to certain types of cancer: *VHL* in renal cancer, for example, and *BRCA1* in breast and ovarian cancer. A more extensive form of DNA hypermethylation occurs in some cases for certain cancers; we discuss this later in the chapter, in the context of links between metabolic and epigenetic dysregulation in cancer.

Genome–epigenome interactions and epigenetic initiation of tumorigenesis

Genetic and epigenetic alterations in cancer used to be regarded as separate mechanisms. Now we know that they work closely together, complementing each

other towards promoting cancer development. As a result, cancer cells can attain greater *plasticity*; by accelerating the normal rate of genetic changes, and promoting dedifferentiation, they can escape normal cellular controls.

Table 10.6 lists some major genome–epigenome interactions in cancer. In addition to genomic changes causing changes to the epigenome, epigenetic changes can provoke major changes to the genome, including increased chromosome instability and silencing of tumor suppressor genes. Because epigenetic changes can be environmentally induced, it is conceivable that epigenetic changes may sometimes initiate tumorigenesis. Take the example of chronic inflammation that can arise out of certain diseases or certain bacterial infections. There is a tight association between chronic inflammatory bowel diseases such as ulcerative colitis and Crohn's disease and the subsequent development of colorectal cancer. And chronic inflammation arising from infection with the bacterium *H. pylori* is also the biggest risk factor for developing gastric cancer. Signaling molecules in inflammatory-associated signaling pathways such as the NF-kB pathway can induce epigenetic changes involved in tumorigenesis.

TABLE 10.6 EXAMPLES OF GENOME-EPIGENOME INTERACTIONS IN CANCER

Genome			**Epigenome**	
Change	Mutation in genes encoding an epigenetic modulator	⇨	Epigenetic dysregulation	Effect
	Mutation in genes encoding an epigenetic modifier*	⇨	Epigenetic dysregulation	
Effect	C→T substitutions	⇦	Deamination of 5-meC**	Change
	Chromosome instability	⇦	Hypomethylation of highly repetitive DNA	
	Silencing of tumor-suppressor genes	⇦	Induced epigenetic changes***	

* See Figure 10.21 for the example of mutation of isocitrate genes that results in dedifferentiation.

** By one estimate, over 60% of the point mutations in the genomes of tumors of internal organs (in tissues that are shielded from UV radiation) arise in CpG sequences.

**** As a result of genetic changes (top two rows) or non-genetic changes, such as altered cell signaling in inflammation, and so on.

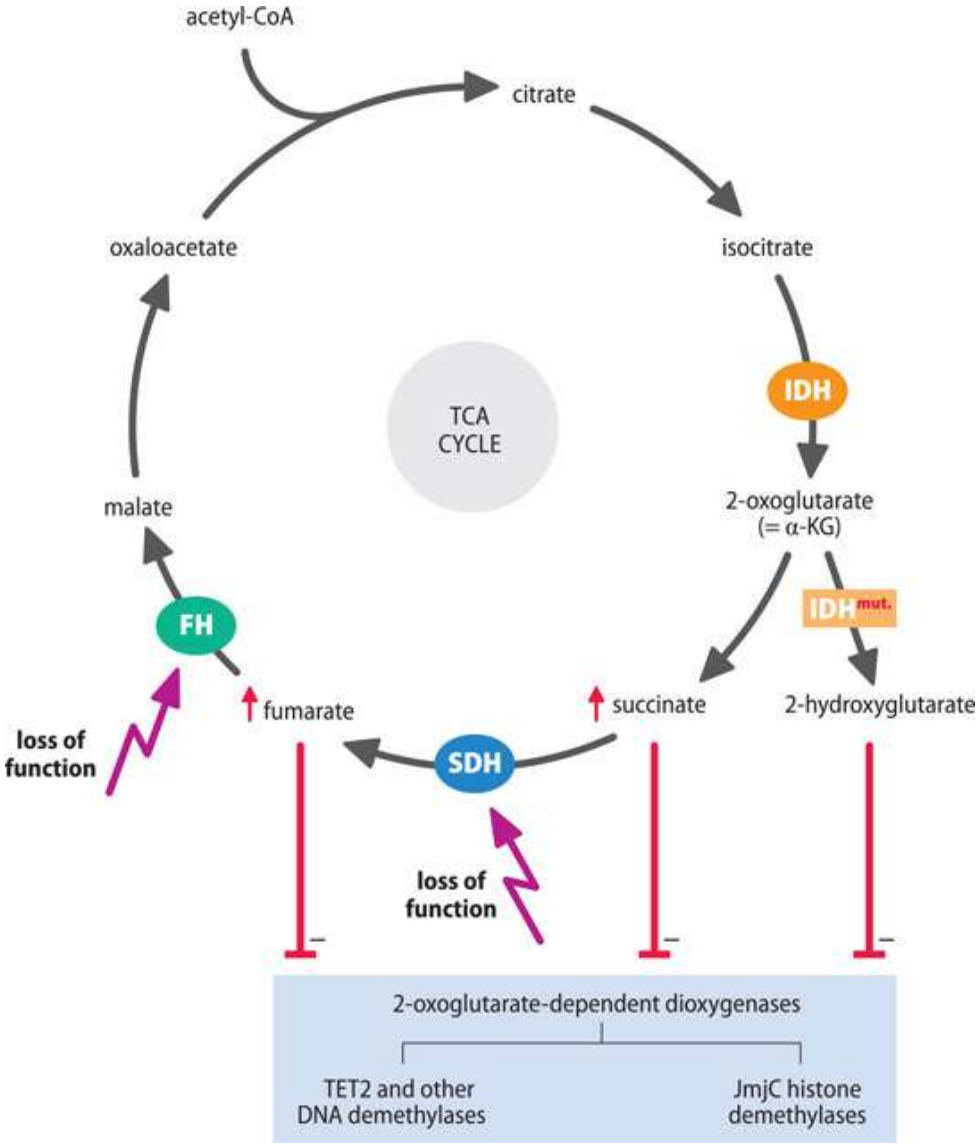


Figure 10.21 Mutation of certain genes encoding enzymes of the tricarboxylic acid (TCA) cycle can cause epigenetic modifications that contribute to cancer. Normal alleles of the *IDH1* and *IDH2* genes produce an isocitrate dehydrogenase enzyme that converts isocitrate to 2-oxoglutarate (also called a-ketoglutarate; a-KG). In certain cancers, certain missense mutations in *IDH1* (R132H, R132C) or in *IDH2* (R140Q, R172K) result in a mutant isocitrate dehydrogenase ($IDH^{mut.}$) that can convert the 2-oxoglutarate made by a normal IDH allele to 2-hydroxyglutarate. This abnormal oncometabolite changes the epigenetic profile of the cell, reversing differentiation to make the cell more like a stem cell. It does that by inhibiting multiple enzymes that depend on using 2-oxoglutarate as a cofactor, including some DNA demethylases such as TET2 and certain histone demethylases of the JumanjiC (JmjC) class. High levels of succinate and fumarate can also inhibit the 2-oxoglutarate-dependent enzymes, as when two loss-

of-function alleles result in a genetic deficiency of fumarate dehydrogenase (FH) or succinate dehydrogenase (SDH), causing a buildup of substrate (red arrows).

10.4 NEW INSIGHTS FROM GENOME-WIDE STUDIES OF CANCERS

Until quite recently, molecular genetic studies of cancer cells had focused on individual genes of interest. Databases were established to store information on DNA changes associated with cancer-associated changes in important cancer-susceptibility genes, such as the International Agency for Research on Cancer's *TP53* database (http://p53.iarc.fr/).

Once the sequence of the human (euchromatic) genome had been obtained, the age of *cancer genomics* could begin. Different genome-wide screens were devised to get comprehensive data from cancer cells, beginning with microarray studies that reported the relative abundance of transcripts from thousands of different human genes. To seek out novel cancer-susceptibility genes, whole genome association studies of the kind described in Section 8.2 have been used, and have been useful (interested readers can find an example at PMID 32424353), but whole exome and whole genome sequencing have been especially fruitful. More recently, high-resolution genome-wide DNA methylation screens have been carried out for some types of cancer, and multiple single-cell genomics and transcriptomic studies have been carried out.

After the launch of the Cancer Genome Project in the UK in 2000, and The Cancer Gene Atlas (TCGA) in the USA in 2006, the International Cancer Genome Consortium (ICGC) was created in 2007 to coordinate efforts on a global scale. The burgeoning data coming out of the cancer genome projects are stored in dedicated databases and can be navigated with dedicated Web browsers (**Table 10.7**). In addition to transforming our understanding of cancer—we describe below some examples of new insights that have emerged—the new data will also have important consequences for cancer diagnosis and treatment.

TABLE 10.7 EXAMPLES OF DATABASES, WEB BROWSERS, AND NETWORKS IN CANCER GENOMICS

Electronic resource	Description	Website URL

Electronic resource	Description	Website URL
COSMIC* database	stores and displays somatic mutation information and related details and contains information relating to human cancers; includes a census of human cancer genes at http://cancer.sanger.ac.uk/census/	http://cancer.sanger.ac.uk/cosmic/
International Cancer Genome Consortium (ICGC)	applications can be made to access controlled data from the ICGH whose goal is to get a comprehensive description of genomic, transcriptomic, and epigenomic changes in multiple different tumor types and/or subtypes that are of clinical and societal importance across the globe	https://daco.icgc.org
The Cancer Genome Atlas (TCGA)	cancer genomics network of research centers in the US	http://cancergenome.nih.gov

* Catalog of Somatic Mutations in Cancer

Genome sequencing has revealed extraordinary mutational diversity in tumors and insights into cancer evolution

Massively parallel DNA sequencing (also called *next-generation sequencing*) is transforming genetics because it offers a huge step up in DNA sequencing output. Because of the extraordinary complexity of cancer evolution it has been extensively applied to sequencing cancer genomes.

Initially, exome sequencing was used to analyze the DNA of cancer cells (the great majority of cancer genes make proteins, and many of the mutations occur within exons). Exome sequencing cannot readily detect copy number variation,

and so cancer exome sequencing projects have been supplemented by genome-wide screens for copy number variation. More recently, sequencing of whole (euchromatic) cancer genomes has also been carried out to reveal all classes of change in somatic DNA in a tumor (when referenced against a corresponding normal tissue genome from the individual).

The first whole cancer genome to be sequenced—an acute myeloid leukemia genome—was reported in 2008. Since then, large numbers of whole cancer genome sequences have been determined. Using cloud computing, the Pan-Cancer Analysis of Whole Genomes (PCAWG) Consortium has facilitated international data sharing. In 2020 it reported the analysis of 2658 whole-cancer genomes and matching normal tissues across 38 tumor types, representing most common cancers.

The extraordinary volume of data pouring out of cancer genome sequencing is collated in different databases, notably the COSMIC database (see Table 10.7). Working out what the huge amount of sequence data means is inevitably a challenge, but already some valuable insights have been revealed.

Mutation number

How many mutations are there in a cancer? There are certainly more than we used to think. From multiple sequenced cancer genomes we now know that adult cancers often have between 1000 and 10 000 somatic substitutions across the genome. However, some types of cancer—medulloblastomas, testicular germ cell tumors, and acute leukemias, for example—have relatively few mutations; others, such as lung cancers and melanomas, have many more mutations (sometimes more than 100 000). If we focus on just the coding sequence (1.2 % of the genome) and consider only the nonsynonymous mutations (which, by changing an amino acid, are more likely to have an effect on cell function than, say, synonymous mutations), the number of nonsynonymous mutations per tumor continues to show a clear dependence on the type of tumor (**Figure 10.18**).

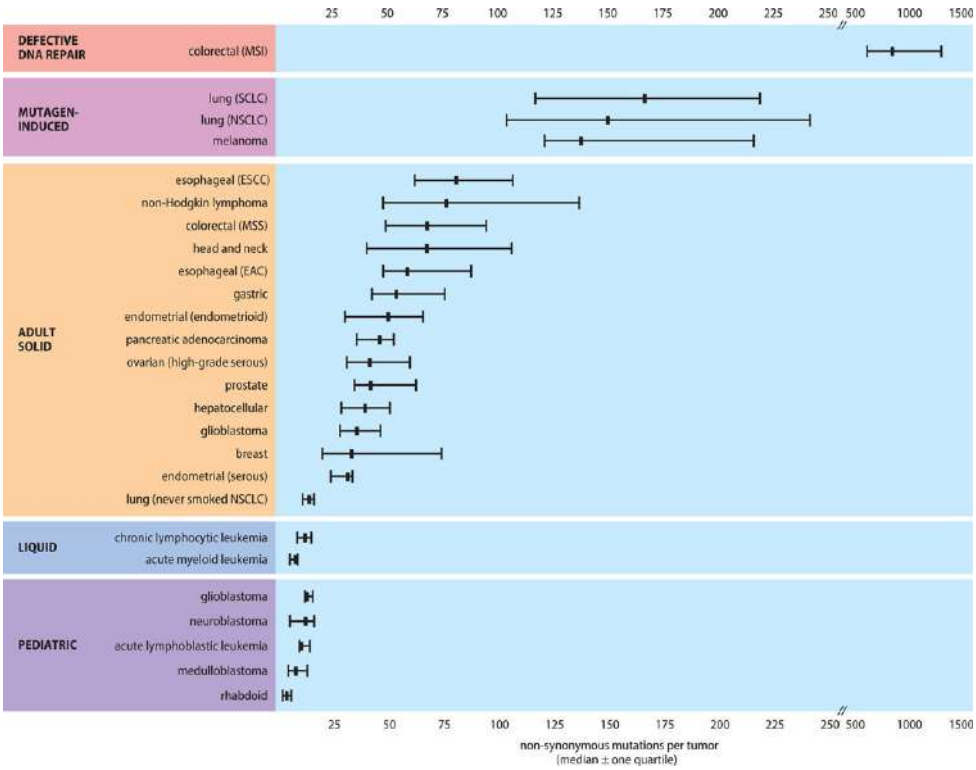


Figure 10.18 Variation in the number of somatic nonsynonymous mutations per tumor in representative human cancers. The median number of nonsynonymous mutations per tumor is estimated from genome-wide sequencing in tumors. Horizontal bars indicate the 25% and 75% quartiles. MSI, microsatellite instability; SCLC, small-cell lung cancers; NSCLC, non-small-cell lung cancers; ESCC, esophageal squamous cell carcinomas; MSS, microsatellite stable; EAC, esophageal adenocarcinomas. (Data from Vogelstein B et al. [2013] *Science* 339:1546–1558; PMID 23539594.)

How can we explain the differences in mutation number? In part this is because different cancers can vary in the number of cell divisions separating the fertilized egg and the cancer cell. And differences in mutation rate at the cell divisions from the fertilized egg cell to the cancer cell must be a factor. Tumors in children or young adults might have lower mutation prevalence simply because the cancer cell has been through comparatively few mitoses. The high mutation prevalence in lung cancers and melanomas most probably reflects an exceptionally high exposure or vulnerability to specific mutagens (tobacco carcinogens and UV radiation, respectively). Inevitably, perhaps, the highest mutation frequencies are found in cancers having a mutation that causes a defect in mismatch repair; see Figure 10.18).

Mutational processes and cancer evolution

Cancer genome and exome sequencing has permitted comprehensive studies of the mutational processes involved in the evolution of a cancer. In 2012 a series of papers provided the first comprehensive dissection of breast cancer. To the cancer surgeon, one of the most striking things about breast cancer is how it progresses differently in each patient, and how each patient responds differently to therapy. This is where molecular genetics might make a difference by helping identify subclasses of tumors with distinct properties that allow different treatment options to be applied, depending on the tumor subtype. The breast cancer studies revealed extraordinary mutational diversity, with multiple independent *mutational signatures*; they also indicated that in most such cancers more than one mutational process has been operative.

Specific mutational signatures found in some cancers simply reflect excessive exposure to specific environmental mutagens that preferentially cause particular types of mutation (for example, UV radiation causes preferential C:G ® T:A transitions in melanoma, and tobacco carcinogens cause preferential C:G ® A:T transversions in lung cancer). Splicing mutations are particularly common in some types of cancer, notably myelodysplastic syndrome (MDS) and chronic lymphocytic leukemia (CLL). This happens because in these cancers some genes that make components of the RNA splicing machinery (such as *U2AF1* in MDS, and *SF3B1* in both MDS and CLL) are frequently mutated.

The comprehensive studies of breast cancer were the first to illustrate just how complex the mutational processes are. One novel mutation process initially discovered from sequencing breast cancer genomes is a type of hypermutation called *kataegis* (from the Greek word for thunderstorm). If a breast cancer genome shows, say, 10 000 tumor-specific mutations, one might expect that the mutations would be mostly randomly distributed across the genome. In that case the average density of the mutations would be 10 000 per 3 Gb or roughly 1 mutation in every 300 kb of DNA. But sometimes highly clustered mutations of the same type are seen, such as C ® -T mutations (**Figure 10.19**). This type of hypermutation appears to result from excess activity of cellular APOBEC proteins (which naturally act as cytidine deaminases in processes such as antibody diversification and RNA editing). By promoting excess activity by these enzymes, tumors find yet another way of generating multiple mutations that natural selection can work on to promote cancer development.

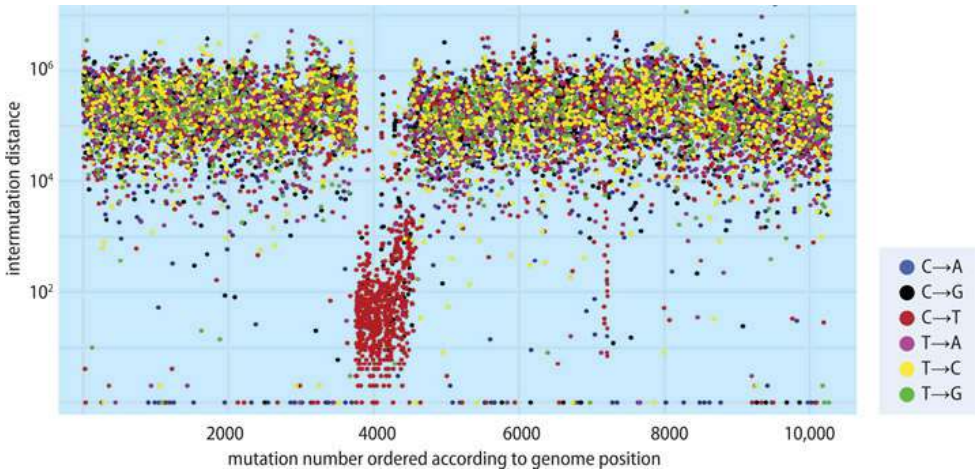


Figure 10.19 A rainfall plot showing an example of kataegis, a form of clustered hypermutation, in breast cancer. Cancer-specific mutations (a total of more than 10 000 from across the genome in this case) are ordered on the horizontal axis according to their position in the genome, starting from the first variant on the short arm of chromosome 1 (at the extreme left) to the last variant on the long arm of chromosome X (position number about 10 500) and are colored according to mutation type (see the color key at the right). The vertical axis shows the distance between each mutation and the one before it (the intermutation distance), plotted on a logarithmic scale. Most mutations in this genome have an intermutation distance of about 10^5 bp to about 10^6 bp, but the plot clearly shows a major region of hypermutation roughly centered on mutation position 4000 (corresponding to a 14 Mb region on the long arm of chromosome 6), where there is an extraordinary clustering of C ® T mutations (red dots) that are spaced from their nearest neighbors by very short distances (often 100 bp or less). Within this region there are defined very short regions of intense C ® T mutation clustering as shown in Figure 4 of the original article. (From Nik-Zainal S et al. [2012] *Cell* 149:979–993; PMID 22608084. With permission from Elsevier.)

Intertumor and intratumor heterogeneity

Genome sequencing has provided the first full understanding of mutational differences between different tumors of the same type, and of mutational differences within tumors. Most differences are due to passenger mutations. When coding sequences were scanned, driver mutations were found to be common in key cancer-susceptibility genes. For example, mutations in *APC, TP53*, and *KRAS* were found to be frequent in colorectal tumors, and the *BRAF* gene was implicated in more than 60 % of melanoma tumors.

The first comprehensive insights into intratumor heterogeneity came from a study of renal cancer by Gerlinger et al. in 2012 (PMID 22397650) in which exome sequencing analyses were carried out on multiple biopsies taken from a primary renal carcinoma and associated metastases from a single individual. Only about one-third of the 128 somatic mutations detected by exome sequencing in the different biopsies from this individual were present in all regions and only one driver gene—*VHL,* the von Hippel–Lindau tumor suppressor gene—was mutated in all analyzed regions. Another driver gene, *SETD2* (which encodes the histone H3K36 methyl-transferase), showed three distinct mutations associated with different regions, and selection pressure was inferred to have found three different ways of inactivating the *SETD2* gene to produce similar tumor phenotypes.

In the above study, in addition to the differences in mutation profiles between biopsies from tumors from different regions, even a single biopsy appeared to consist of two different clonal populations. The cells that seeded metastases seem to have diverged at an early stage from those that formed the primary tumor; the two groups had a differentiating series of mutations, and the cells that would form metastases lacked a mutation in a driver gene, *MTOR* (which encodes the mammalian target of rapamycin kinase).

Defining the landscape of driver mutations in cancer and establishing a complete inventory of cancer-susceptibility genes

As described below, proteins made by known cancer genes have become targets for successful anti-cancer drug development. Identifying new cancer-susceptibility genes has therefore been a key goal of cancer genome studies. Until quite recently, most known cancer genes had been identified by three approaches: analyses of associated chromosome abnormalities (notably translocations) in which breakpoints could be identified by FISH; candidate gene studies (using information from experimental model organisms); and studies of copy number variation (by detecting oncogene amplification or by scanning for loss of heterozygosity). Linkage analyses had also been important in defining some genes that underlie inherited cancers, such as the *BRCA1* and *BRCA2* genes that are important susceptibility genes in breast and ovarian cancer.

To identify novel cancer-susceptibility genes, genomewide association (GWA) studies were initially employed, but they have had limited success. Instead, genome or exome sequencing of multiple tumors has become the preferred

approach: the sequences are referenced against normal somatic cells from the relevant individuals to identify tumor-specific mutations.

How can genome-wide sets of tumor-specific mutations allow us to identify cancer-susceptibility genes? The driver mutations for any type of cancer might be expected to be confined to a comparatively small set of key cancer genes that are frequently mutated in that type of cancer (and which are presumably crucial to its evolution). Passenger mutations, by contrast, might be expected to be rather randomly distributed across the genome and to be somewhat different in unrelated tumors of the same cancer type. By looking at multiple tumors of the same type, one might expect to quickly discriminate between driver and passenger mutations. (However, it is not always so straightforward: a cluster of somatic mutations may also be attributable to an increased local mutation rate; in that case, passenger mutations may initially be confused with driver mutations.)

Cancer gene and driver mutation distribution

Genome-wide sequencing approaches have been enormously successful, not just in identifying novel cancer-susceptibility genes but also in defining the distribution of the cancer-susceptibility genes in different cancers and the profile of associated driver mutations. In a study of 100 breast cancer tumors, for example, a total of 250 driver mutations were found; the number of driver mutations per tumor ranged up to a maximum of six, with an average of 2.5 (**Figure 10.20**).

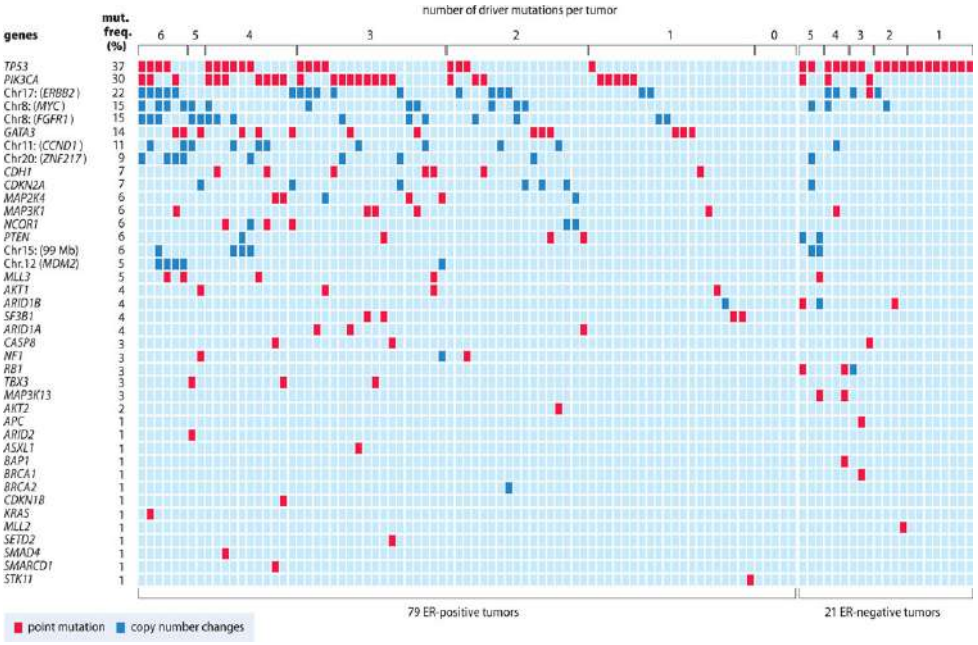


Figure 10.20 The landscape of driver mutations in a study of 100 primary breast cancers. Of the 100 cancers, 79 expressed estrogen receptor (ER-positive), and 21 were ER-negative. By referencing against control DNA samples from normal cells, somatic mutations were identified in 40 cancer-susceptibility genes (shown in the left-hand column). Point mutations were identified by whole exome sequencing. Changes in copy number (shown in blue) include amplification of oncogenes and loss of alleles in tumor suppressors; they were identified by hybridization to whole genome SNP (single nucleotide polymorphism) arrays (described at http://www.sanger.ac.uk/genetics/CGP/CopyNumberMapping/Affy_SNP6.shtml). At least one of these genes or loci was mutated in all of the tumors, except in five ER-positive tumors. A maximum of six of the genes were mutated in any one tumor. The most significant (frequent) cancer genes were found to be *TP53* (mutated in 37 % of all tumors, and in close to 90 % of ER-negative tumors) and *PIK3CA* (mutated in 30 % of the tumors). Mut. Freq., mutation frequency. (Adapted from Stephens PJ et al. [2012] *Nature* 486:400–404; PMID 22722201. With permission from Macmillan Publishers Ltd.)

In the above study, seven genes—*TP53, PIK3CA, ERBB2, MYC, FGFR1/ZNF703, GATA3,* and *CCND1*—were found to be mutated in 10 % or more of tumors, and collectively these genes were the source of almost 60 % of the driver mutations detected in coding sequences (see Figure 10.20). In a parallel exome sequencing study of 510 breast cancers, also published in 2012, *TP53* and *PIK3CA* were also found to be the most frequently mutated genes. More recently, whole genome analyses have extended the hunt for driver mutations into

noncoding DNA. Although some nonocoding driver mutations are common (such as in the *TERT* promoter to drive expression of telomerase), data from the Pan-Cancer Analysis of Whole Genomes suggest that noncoding driver mutations are rare compared to driver mutations in coding DNA, possibly as a result of a lack of discovery power.

Novel cancer-susceptibility genes

The genome-wide sequencing projects have recently delivered many novel cancer-susceptibility genes. However, the genes being discovered are ones that are infrequently mutated; as with other complex diseases, the major cancer-susceptibility genes have previously been identified. It may be that there will be a considerable tail of low-frequency cancer-susceptibility genes.

By March 2022 the Cancer Gene Census within the Cosmic database (at http://cancer.sanger.ac.uk/census/) had listed a total of 578 identified human cancer susceptibility genes. Until recently, the focus has very much been on looking at coding sequences. However, a study of regulatory regions in melanomas has emphasized the need to look at noncoding regions: whole genome sequencing found that highly recurrent somatic mutations occur at two specific nucleotides in the promoter of the telomerase reverse transcriptase gene, *TERT*. Further studies show that the effect of the mutations is to generate a binding site for the ETS transcription factor that upregulates *TERT* expression. The *TERT* promoter mutations were found to occur in more than 70 % of melanomas and about one in six of the other types of tumor examined in the study.

Novel cancer-susceptibility genes are going to be drawn from the genes that support the different biological capabilities of cancers. As we explain in the next subsection, many might not be the conventional oncogenes and tumor suppressor genes that we have become familiar with.

Non-classical cancer genes linking metabolism to the epigenome

One of the surprises emerging from cancer genome sequencing has been the extent to which genes that work in metabolism are important in cancer. These genes can

be non-classical oncogenes or tumor suppressor genes, and many of them have been linked to epigenetic regulation.

Take, for example, the *IDH1* and *1DH2* genes that make respectively cytosolic and mitochondrial isocitrate dehydrogenase, enzymes that work in the tricarboxylic acid (Krebs) cycle to convert isocitrate to 2-oxoglutarate (also known as a-ketoglutarate). One of these two genes is (heterozygously) mutated in 80–90 % of adult grade II/III gliomas and secondary glioblastoma, in more than 50 % of chondrosarcomas, in a significant proportion of acute myeloid leukemias, and in some other cancers. In terms of mutation types and distribution, the genes clearly fall in the oncogene camp (as previously noted for *IDH1* in Figure 10.9).

The predominant *IDH1/IDH2* cancer-associated mutations are specific missense mutations producing mutant enzymes that convert 2- oxo-glutarate (produced by the normal allele) to 2-hydroxyglutarate. At high concentrations, 2-hydroxyglutarate inhibits multiple enzymes that depend on 2-oxoglutarate as a cofactor and work in epigenetic modification, including certain DNA demethylases, such as TET2, and various histone demethylases. That can cause reprogramming of the cell to make it less differentiated (**Figure 10.21**).

As well as oncogenes, tumor suppressor genes regulate the epigenetic-metabolic link in cancer cells. The paper published by Sebastian et al. in 2012 (PMID 23217706) gives the example of the SIRT6 tumor suppressor, a histone deacetylase that normally suppresses aerobic glycolysis.

Tracing the mutational history of cancers: just one of the diverse applications of single-cell genomics and transcriptomics in cancer

The genomics revolution has recently produced an extraordinary technological achievement: the capacity to carry out simultaneous single-cell analyses of genomes and transcriptomes in populations of cells. Not unexpectedly, given the capacity for single body cells to mutate and evolve into rogue cancer cells, applications in oncology have been right at the forefront. **Figure 10.22** illustrates the diverse applications in cancer research.

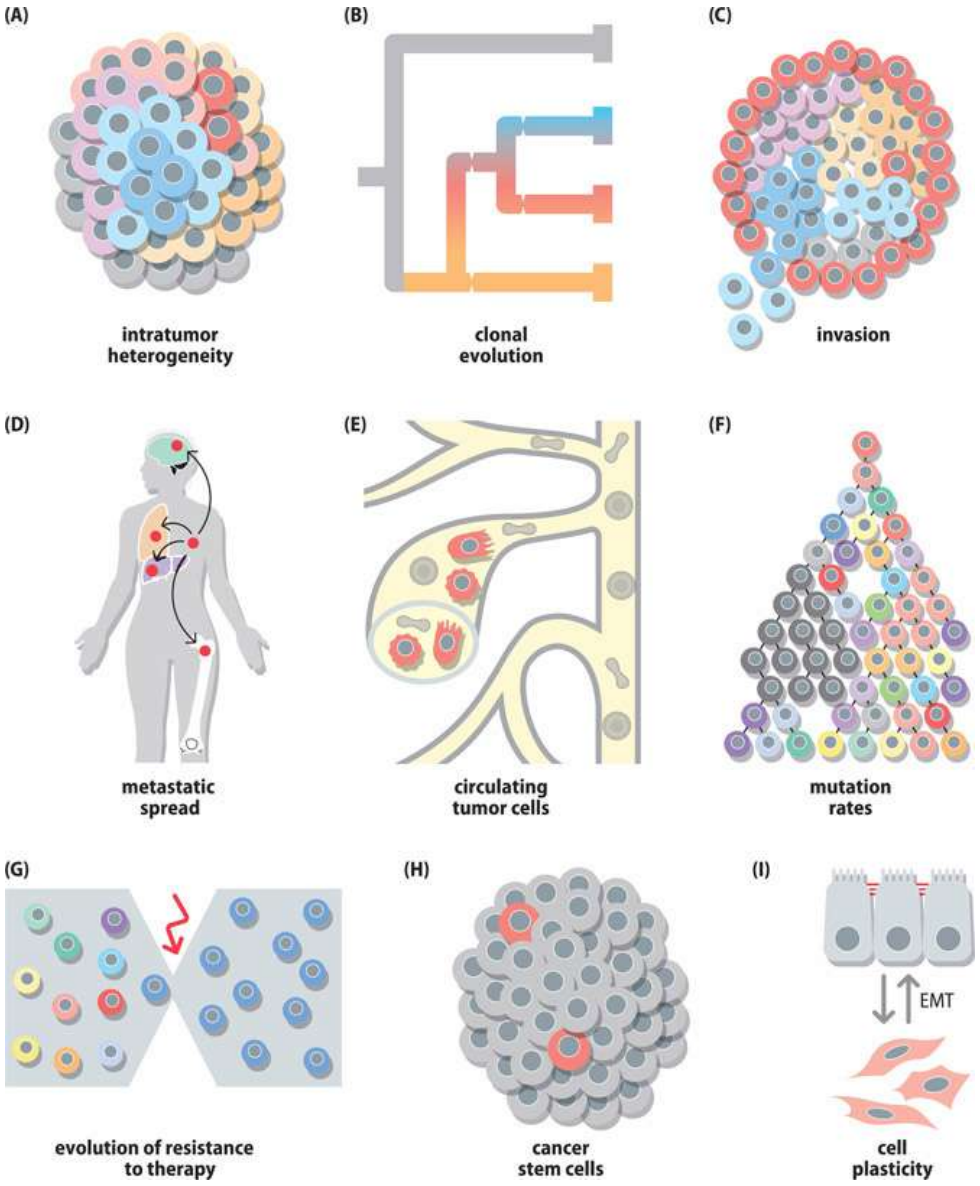


Figure 10.22 Applications of single-cell genomics and transcriptomics in cancer research. EMT, epithelial-mesenchymal transition. [Adapted from Lindau et al. (2015) *Nature* 526:525–530; PMID 26466571]. With permission from Springer Nature Copyright© 2015.

As an illustration of the power of single-cell genomics, consider how it is applied to study tumor evolution. One can always get an idea of the general evolution of a tumor class by comparing early-stage tumors with pre-cancerous lesions and late-stage tumors. But there is the problem of heterogeneity within tumors: a tumor is not simply a clonal colony of cells. Rather, because increasing genomic instability drives accelerated mutation, a tumor is a heterogeneous

population of cells related by mutational branchpoints of the type shown in Figure 10.22B.

By carrying out single-cell genomic DNA sequencing in tumors, we can position cells of the tumor in phylogenetic lineages to reveal the nature and order of successive driver mutations. Since the mutational landscape in leukemia is relatively simple (in comparison with that of most solid tumors), a single blood sample is enough to allow the evolution of a tumor to be reconstructed. If driver A is found in all leukemic cells and driver B is present in some cells only, for example, one can infer that mutation A appeared first and subsequently mutation B arose and was transmitted to a subpopulation. For an example of tracing the mutational history of cancers, see **Figure 10.23**. Single-cell transcriptomics in cancer sequencing is also used in various ways to illuminate cancer processes, and we consider them against the broader background of cancer transcriptomics in the next section.

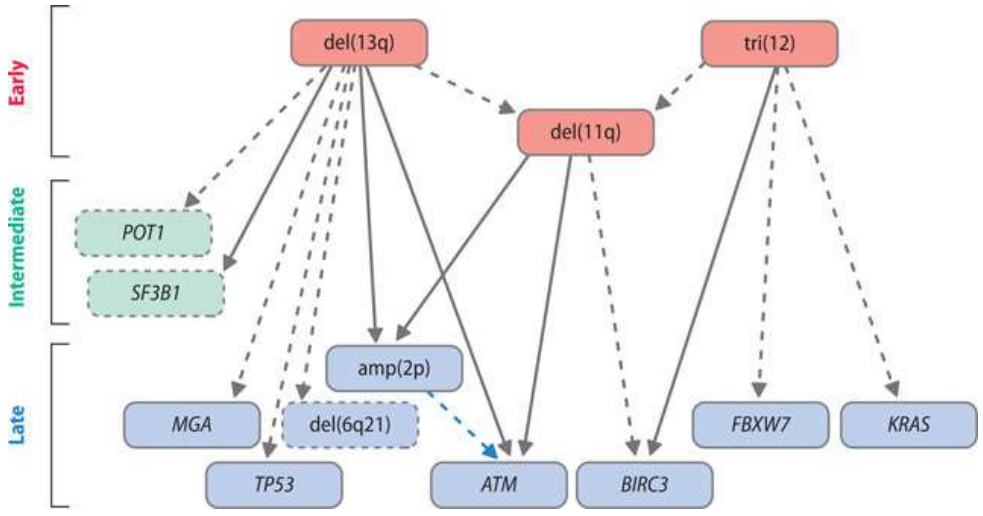


Figure 10.23 Mutational history of chronic lymphocytic leukemia inferred from the clonality of mutations after single-cell sequencing analyses on a single blood sample. Early events are a few types of large copy number changes due to chromosomal instability, but there is substantial diversity in the late driver mutations. (Reprinted from Landau et al. [2015] *Nature* 526:525–530; PMID 26466571.)

Genome-wide RNA sequencing enables insights into the link between cancer genomes and cancer biology and aids tumor classification

Following on from the huge effort in cancer genome analysis, genome-wide analyses have been carried out to study how genetic changes in tumor cells are expressed to alter the biology of the cells. Most of the effort has been expended on studying RNA transcripts from cancer cells. The first genome-wide studies of cancer transcriptomes used microarray-based expression analyses, but modern analyses use **RNA-Seq** (RNA transcripts are first converted to DNA using reverse transcriptase, then sequenced).

The interest in broadening genome-wide studies of cancer to include additional "omics", notably transcriptomics and proteomics, has been driven by two major aims. The first is to be able to link genome changes to cancer biology. In this regard, the Paull et al. (2021) reference under Further Reading has been a major advance, linking alterations to the genome to the transcriptional profiles of diverse types of cancer cells. It identified 407 proteins as master regulators that, by working together in groups with variable modular combinations, convert the genomic changes into 112 transcriptionally distinct tumor subtypes.

The second major aim has been to develop some molecular classification of tumors that might improve on conventional methods (which largely depend on how tumors appear when pathologists examine them under the microscope). Improved classification of tumors into subgroups has the potential for more efficiently targeting clinical actions such as prognosis and treatment. New molecular classification methods have recently become available for some types of cancer, such as a study reported in 2021 that identified four molecular subtypes of small cell lung cancer (PMID 33482121). For other cancers such as breast cancer, previous knowledge has been built upon using new molecular subtyping arising from various genome-wide "omics" technologies as reported in the Parsons and Francavilla (2020) review under Further Reading. This area is an evolving one, and clinical benefit from molecular subtyping will be progressively accrued.

The need to focus on biological pathways important in cancer cell evolution

If we view the genome changes in cancers to be complex, linking them in logical ways to diverse changes in the transcriptome and proteome takes the level of complexity to new levels, and will keep researchers busy for quite some time.

To devise clinical benefit from all these studies, it may be productive to concentrate on the effects of the genetic changes within cancer cells. The huge complexity of all the individual mutations that contribute to cancer can be reduced if we focus not just on the (still very heterogeneous) collection of individual cancer susceptibility genes, but on how key regulators work in biological pathways that are important in cancer cells. Whatever altered genomic states and altered transcriptional states arise as cancers evolve, the functional endpoint is signifi-cant changes in key proteins working in various cell signaling pathways (there is little evidence of driver mutations in genes that make noncoding RNAs). In an overview of molecular events in cancer reported in 2013, Bert Volgelstein and colleagues identified 64 high-penetrance oncogenes and 74 high-penetrance tumor suppressor genes, but reported that they all act through one or more of just 12 cell signaling pathways (see **Figure 20.24A**).

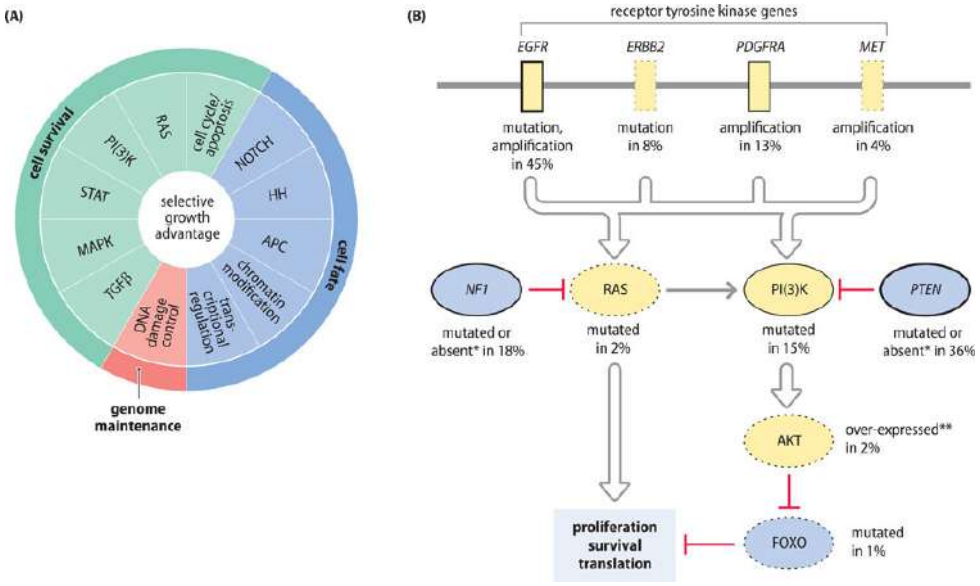


Figure 10.24 Reducing complexity by focusing on signaling pathways important in cancer. (A) Twelve cancer cell signaling pathways and the processes they regulate. All of the pathways confer a selective growth advantage. The pathways can be organized into three core cellular processes (outer ring). Note that *TP53* encoding the p53 master regulator functions in both the genome maintenance and cell cycle/apoptosis pathways. From Vogelstein B et al. [2013] *Science* 339:1546–1558; PMID 23539594. Reprinted with permission from the AAAS. (B) The importance of the RAS/PI(3)K pathway in glioblastoma multiforme. Oncogene products are shown in yellow shading; tumor suppressor proteins are in blue shading. Red T-bars indicate inhibition. The four upstream receptor typrosine kinase genes and the genes making each of the six downstream proteins are all mutated, to different extents, in glioblastoma multiforme tumors.

*Via homozygous gene deletion. **Via gene amplification. Data from the Cancer Genome Atlas Research Network [2008] *Nature* 455:1061–1068; PMID 18772890.

For some cancers, what seems a hugely complex series of associated mutational changes and even of genes conferring susceptibility to that cancer, the picture may be altogether simpler at the level of cell signaling pathways. Take the example of glioblastoma multiforme where tumors show high levels of genetic heterogeneity. Despite the large number of genes involved, they all work in two major signaling pathways, a RAS-PI(3)IK pathway (see Figure 10.24B) plus the cell cycle-apoptosis pathway.

Single-cell transcriptomics

Classification of tumor cell diversity and the ability to identify rare tumor cell populations is necessarily limited when analyzing bulk tumor cell populations. The development of droplet-based microfluidic single-cell RNA-Seq enables high-throughput capture and molecular barcoding of individual cells that can be analyzed rapidly. This field is still young, but the review published by Kim et al. in 2020under Further Reading provides an example of the kinds of applications that are being found.

10.5 GENETIC INROADS INTO CANCER THERAPY

As described in Section 8.3, complex diseases are caused by a combination of genetic and environmental factors, and cancers are no different in this respect. Before we go on to look at therapeutic approaches directed at genetic control points in cancer, it is important to acknowledge the huge effect of environmental factors in cancer. In addition to well-established connections between UV radiation and melanoma, and between tobacco carcinogens and lung cancer, many other cancers are strongly determined by environmental factors. Rates of colon cancer, for example, vary as much as 20-fold between countries; the dramatic differences are due to environmental factors, specifically dietary components, rather than genetic susceptibility. That is evident when a population migrating from one country to another exhibits a colon cancer rate typical of the new country

within one or two generations of settling there. Microbial infections play their part too, and not just viral infections associated with cancers but also certain chronic bacterial infections. The outstanding example is *Helicobacter pylori*, a gastric pathogen that colonizes ~50 % of people across the world and persists for the lifetime of the host. Infection with *H. pylori* causes chronic inflammation and is the strongest known risk factor for gastric cancer, the second most frequent cause of cancer-related deaths worldwide.

How can the burgeoning knowledge of the underlying genetics of cancer have a clinical impact? As previous sections of this chapter testify, the revolution in cancer genomics has made clear the complexity of cancer evolution, and also the extraordinary degree of both intratumor and intertumor heterogeneity. This can pose difficulties in validating biomarkers for the oncogenic process: biopsies from the same tumor may show different genetic profiles. And, because of intra-tumor heterogeneity, natural selection can be expected to propel the growth of drug-resistant clones.

Treatment or prevention?

Faced with these problems, should we simply accept that treating cancers is never going to be anything more than damage limitation, disease management rather than cure? Maybe, but as described below, there may yet be grounds for optimism. But, undeniably, genetics—and especially genomics—has shone a bright torch into the gloom that used to shroud the inner workings of many cancers. The result is a much more informed understanding of the fine, granular detail of the underlying mutation mechanisms, a greater appreciation of the molecular characteristics of cancers, and detailed insights into how cancers evolve. Once we have fully understood the molecular pathways of cancer and cancer evolution in fine detail, we may be in a much better position to devise novel treatments.

Targeted anticancer therapies are directed against key cancer cell proteins involved in oncogenesis or in escaping immunosurveillance

Traditional cancer treatments have been blunt tools: surgery to excise tumors and chemotherapy or radiotherapy to kill them. The latter two methods are simply designed to kill actively dividing cells; the problem, of course, is that they also kill actively dividing normal cells, adversely affecting the health of the patient.

Despite their limitations, the long-established triad of blunt-tool methods are still frequently used in cancer treatment today.

Genetic and especially genomic analyses have recently identified many cancer-susceptibility genes, enabling multiple opportunities to developed *targeted* anti-cancer therapies, directed at specific proteins directly or indirectly involved in oncogenesis. Good targets should be present in cancer cells but not normal ones, or should be strongly upregulated in cancer cells when compared to normal cells. As described in the subsections below, targeted anticancer therapies involve designing molecules that specifically bind to and inhibit key proteins involved in oncogenesis, or that can selectively kill cancer cells (either directly, or indirectly through the intervention of T cells). We cover targeted therapies using small molecular drugs or monoclonal antibodies here. A third method, using genetically engineered T cells, is described in the section following the next one.

Targeted therapies using small molecule drugs

Conventional small molecule drugs produced by the pharmaceutical industry can be screened for evidence of binding to a specific protein of interest. Often a molecule like this can fit snugly in a cleft in the protein, sometimes disrupting the function of the protein.

The first successful targeted anticancer therapy was achieved using this approach more than three decades ago. It was prompted by the *BCR-ABL* chimeric oncogene on the Philadelphia chromosome frequently found in chronic myeloid leukemia (CML), as shown in Figure 10.8A. The rationale was this: if the *BCR-ABL* gene is present in tumor cells but not in normal cells, can we find a drug to specifically bind to the BCR-ABL fusion protein and stop it working? And if so, shouldn't the drug selectively stop the tumor cells proliferating? The answer to both questions was yes. Imatinib (marketed as Gleevec^{R}) obtained FDA approval in 2001 to treat CML patients with the Philadelphia chromosome and was seen to be a resounding success. Subsequently there has been a variety of other successes with small-molecule drugs. Some of these have been targeted to bind key proteins implicated in different cancers—see **Table 10.8**.

TABLE 10.8 EXAMPLES OF TARGETED CANCER THERAPIES

Cancer	Drug/MAb	Protein target	Mode of action

Cancer	Drug/MAb	Protein target	Mode of action
SMALL MOLECULE DRUGS			
Breast	Tamoxifen	Estrogen receptor in ER-positive breast cancers	Blocks ER, preventing growth signals
Leukemia, (AML, Ph+)	Imatinib	BCR-ABL1 fusion protein	Inhibits abnormal signaling by fusion protein tyrosine kinase
Leukemia (AML, CLL)	Venetoclax	BCL2, a key inhibitor of apoptosis	Binds to BCL2 to disrupt its function, thereby stimulating apoptosis
Lymphoma (SLL)			
Melanoma	Verumafenib	BRAF V600E mutant protein	Specifically inhibits V600E mutant BRAF, triggering apoptosis
Non-small cell lung cancer	Crizotinib	EML4-ALK fusion protein	Inhibits abnormal signaling by fusion protein tyrosine kinase
Ovarian (advanced, BRCA1/2 minus)	Olaparib*	PARP1 enzyme	Blocks repair of DNA breaks in BRCA1 -mutant cancers
Various (advanced)	Becacizumab (Avastin)	VEGF (vascular endothelial growth factor)	Inhibits angiogenesis
MONOCLONAL ANTIBODIES			
Breast	Trastuzumab	EGF receptor (EGFR) on HER2-positive cells	Attaches to receptor; identifies the cell as a target for the immune system
Leukemia	Rituximab	CD20 B-cell surface protein	Binds to CD20; identifies cells as targets for NK cells.

Cancer	Drug/MAb	Protein target	Mode of action
Lymphoma	Ibritumomab**	CD20 B-cell surface protein	Binds to CD20; carries a radioactive payload to kill cells it binds to
Melanoma	Ipilimumab	CTLA4T-cell inhibitor	Blockade of CTLA4 or PD1 allows T cells to renew attack on cancer cells
	Nivolumab	PD1 T-cell inhibitor	
Prostate (advanced)	Sipuleucel-T***	Prostatic acid phosphatase (PAP)	Stimulates T-cell response against PAP

AML, acute myeloid leukemia. CLL, chronic lymphocytic leukemia. EGF, epidermal growth factor. SLL, small lymphocytic lymphoma. Ph+, Philadelphia chromosome present. PARP, Poly(ADP-Ribose).

* Demonstrates the potential of *synthetic lethality,* where a combination of two nonlethal deficiencies may result in a lethal effect. PARP activates repair of single-strand DNA breaks. When PARP is absent, these breaks have to be repaired by BRCA1/2-mediated homologous recombination, and so when tumor cells lack BRCA1/2, the error-prone nonhomologous end joining method of DNA repair is used, often leading to cell death.

** Ibritumomab is radiolabeled before use by attaching a yttrium 90 radioisioptope.

*** Uses a proprietary protein (PAP fused to GM-CSF to stimulate the patient's own leukocytes *ex vivo.*

Note that the target proteins in some targeted therapies may play a supportive role common to multiple cancer types. For example, to fuel their growth, advanced tumors often develop a vascular supply (Figure 10.5A), providing oxygen and nutrients. Treatment with avastin^{R}, which inhibits *angiogenesis*, the process by which new blood vessels form, inhibits the outgrowth of metastases.

While there have been promising successes for targeted therapies at early stages in treatment, relapses are common as tumors mutate to become resistant to treatment, and we consider this problem below.

Targeted therapies using monoclonal antibodies

Monoclonal antibodies raised against a specific protein target provide an alternative type of drug that can be used as a way of killing cancer cells.

Occasionally, a specific monoclonal antibody, radiolabeled with a cytotoxic radioisotope, permits direct killing of cancer cells—see the example of ibritumomab in Table 10.8. Usually, however, the object is a type of **immunotherapy** to encourage some immune response against cancer cells (see Table 10.8 for examples).

An important example of immunotherapy is *immune checkpoint therapy*. This field was developed independently by Tasuku Honjo and James Allison who shared the Nobel Prize in Physiology or Medicine in 2018 for their work on engineering antibodies to, respectively, the PD1 and CTLA4 cell surface receptors. As part of the self-nonself recognition system, T cells have certain brakes to ensure that they are not inappropriately activated, notably by ligand activation of PD1 and CTL4A. When a T cell encounters a cell with a ligand protein on its surface for the PD1 and/or CTLA4 receptor, the resulting receptor-ligand interaction conveys a signal to inhibit T cell responses. In order to escape immune surveil-lance, cancer cells take advantage by expressing high amounts of the PD1 and/or CTL4A ligands on their cell surfaces. The ipilimumab and nivolumab monoclonal antibodies bind specifically to the PD1 and CTL4A receptors, respectively, blockading them from interaction with ligand proteins, thereby reactivating the capacity of T cells to kill cancer cells.

CAR-T Cell therapy and the use of genetically engineered T cells to treat cancer

In addition to standard small molecule drugs and monoclonal antibody drugs, novel targeted anticancer therapies have been developed using genetically modified T cells, building upon the natural roles of T cells in tumor immunosurveillance. (T cells are able to detect and kill tumor cells arising from both virus infection and from non-viral genetic and epigenetic changes to cells of the body. In the latter case, the tumor antigens with highest specificity are derived from new peptides created through chromosomal translocations and frameshifting mutations.)

Cytotoxic T cells naturally recognize peptide antigens on the surface of cells only after they have been bound by an HLA protein (MHC restriction—see Box 8.3 on page 265), and successful recognition is also dependent on binding of ligands to additional co-receptors on T cells. In the laboratory, T cells can be genetically engineered so that they can bind a specific protein of interest without

the need for an associated HLA protein. To do that cultured T cells are transfected with a gene construct that is designed to make an artificial trans-membrane ***c**himeric**a**ntigen**r**eceptor* (CAR).

The extracellular domain of a T-cell chimeric antigen receptor is designed for antigen recognition. It is typically largely composed of an scFv (single-chain variable fragment) antibody sequence, that is, the variable regions on heavy (VH) and light (VL) chains of a standard monoclonal antibody connected by a short linker peptide. An additional hinge region is purely for structural reasons, enhancing the flexibility of the antigen-binding head of the scFv domain (see **Figure 10.25**).

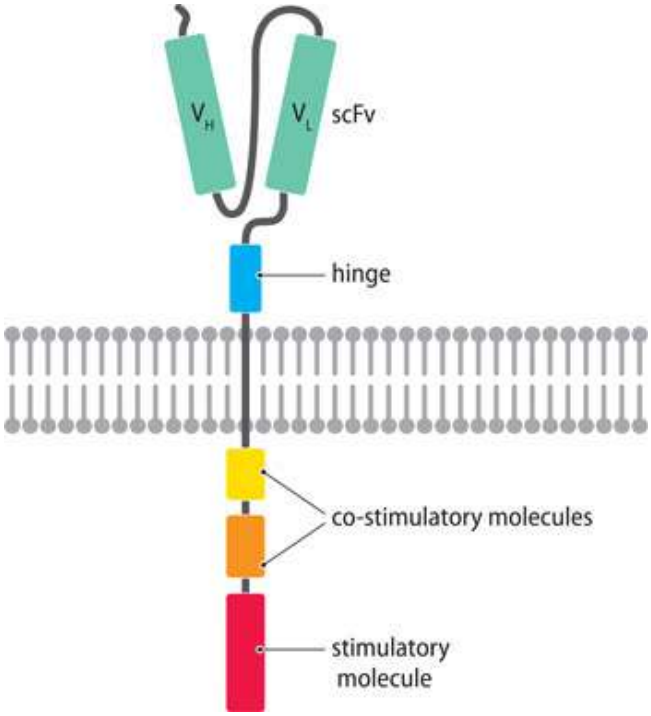


Figure 10.25 Structure of a third-generation chimeric antigen receptor, as deployed in CAR-T cell therapy. The extracellular region is composed of a scFV domain component used in antigen recognition (with the distal ends of the VH and VL variable Ig chains forming the antigen binding site) and a short hinge to allow flexibility. The intracellular region is composed of a stimulatory molecule (often a CD3 zeta chain from a T cell receptor) and two or more co-stimulatory molecules (such as CD27, CD28, OX40, 4-188, and ICOS).

The intracellular region is used in signal transduction. It is formed by bringing together two types of molecule: stimulatory molecules permit downstream activation of the T cells to release cytokines (ultimately leading to death of the

target cell) and interleukins (needed for proliferation), and costimulatory domains serve to enhance the immune response.

Because of *MHC restriction*, recognition of other cells by T cells requires that the cells present antigens bound to HLA proteins, and so cancer cells can switch off HLA expression to escape immunosurveillance by T cells. However, CAR-T cells and other types of genetically engineered T cells remove the requirement for HLA recognition. Treatment with engineered T cells such as these are the equivalent of giving patients a "living drug" that targets a specific protein of choice.

In an early application CAR-T cells were designed to recognize the B-cell antigen CD19 and used to treat B-cell lymphomas and some leukemias with very considerable success. Initially, *autologous* T cells were used: T cells would be removed from a patient then genetically engineered in the lab to make desired CAR-T cells that would be infused back into the patient. For greater convenience and reduced costs, there has been a move towards producing universal, off-the-shelf CAR-T cells; because they would be genetically different to the cells of a patient, there is a requirement for immunosuppressant drugs during treatment.

Despite the very significant clinical benefits, there have been drawbacks in CAR-T cell therapy, notably the propensity to induce "cytokine storms", wherein a massive release of cytokines has resulted in some fatalities in clinical trials. Fine-tuning the protein engineering has become a priority.

The molecular basis of tumor recurrence and the evolution of drug resistance in cancers

Although the initial results of anticancer targeted therapies using small molecule drugs, monoclonal antibodies, or engineered T cells can be very positive, and though they can have significant side effects, the treatments are generally well tolerated when compared to chemotherapy. Applicability has, however, been variable. The therapies have worked well in leukemias, and chronic myelogenous leukemia has been especially amenable given that more than 90 % of patients have the Philadelphia chromosome and produce the BRC-ABL fusion protein that can be treated effectively with Gleevec® (imatinib). But in general, leukemias are amenable to treatment, showing comparatively limited genomic instability and often being identified at an early stage in tumor development. By contrast,

epithelial cancers are not so easily treated, showing much greater genomic instability and not usually being caught until later stages in tumor development.

The basis of tumor recurrence

A general problem in treating cancer is that some cancer cells can survive treatment and tumors frequently recur quite quickly. That poses the question: why? Two possible answers are that cancer stem cells are especially resistant to drug treatment or that subpopulations within a tumor can survive treatment. Support for especially resistant stem cells includes a study on a genetically engineered mouse model of glioblastoma in which a relatively quiescent subset of endogenous glioma cells, with properties resembling cancer stem cells, was found to be responsible for sustaining long-term tumor growth (by producing transient populations of highly proliferative cells). If cancer stem cells are relatively resistant to therapy, they might survive to repopulate a vastly shrunken tumor. If so, the problem becomes how to effectively target and kill populations of cancer stem cells about which we know little.

The possibility of tumor heterogeneity has been amply demonstrated in various studies. If a malignant tumor consists of genetically different populations, some cells might survive drug treatment and natural selection could foster the development of tumor subclones with mutations that render the therapeutic drug ineffective in some way. (There are parallels, therefore, with infectious diseases and the evolution of drug resistance in microbes.)

The evolution of drug resistance

The evolution of drug resistance in targeted cancer therapy can occur in different ways. Sometimes, mutations develop in the gene encoding the drug target itself. For example, in the treatment of chronic myeloid leukemia with imatinib, tumor subclones develop imatinib resistance by developing point mutations that alter the kinase domain of the BCR-ABL1 fusion protein. The mutant kinase retains the catalytic activity required for tumor formation, but its altered structure means that imatinib can no longer bind to it effectively to inhibit it. Drug resistance for many other kinase inhibitors works by a similar mechanism: often the mutations confer

resistance by blocking interactions between drug and target through steric hindrance.

An alternative way of developing drug resistance occurs when the tumor mutates to amplify the drug target gene. Occasionally, for example, resistance to kinase inhibitors in chronic myeloid leukemia is achieved when tumors succeed in amplifying the *BCR–ABL1* gene. Prostate cancers often acquire resistance to drug-mediated androgen deprivation by amplifying the androgen receptor gene.

Yet another option for a tumor to develop drug resistance is to find a way of bypassing the primary drug target (which remains unaltered, and continues to be inhibited by the drug). This can take the form of mutating a downstream effector in the same pathway to render cells insensitive to drug inhibition of a cell surface receptor, for example; or an alternative pathway is activated. For example, the monoclonal antibody trastuzumab is designed to treat breast cancer by binding to and interfering with the human epidermal growth factor receptor 2 (HER2), but tumors can bypass the effects of the drug by activating expression of an alternative receptor, such as HER3.

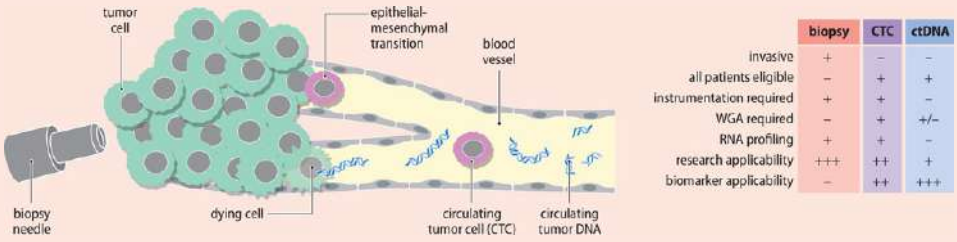
Of course, it would be highly desirable to monitor tumors so as to detect emerging resistance clones as soon as possible to hopefully permit changes in treatment to stop them in their tracks. The recent development of "liquid biopsies" may be an important advance (see **Clinical Box 14**).

CLINICAL BOX 14 LIQUID BIOPSIES IN CANCER: TOWARDS CLINICAL PRACTICE

Diagnosing and monitoring cancer are aided by taking cell samples for examination. While leukemias are conveniently diagnosed and monitored using blood tests, solid tumors have routinely been accessed through invasive biopsies. Most use hollow needles following local anesthetic, but less easily accessed tumors have often required the use of cutting tools attached to an endoscope, or even open or laparoscopic surgery. Monitoring cancer using serial biopsies is generally not a very attractive option, therefore, especially if patients are elderly and frail.

The attractive possibility of **liquid biopsies** in cancer has been made possible by the observation that both circulating tumor cells (CTC) and cell-free

circulating tumor DNA (ctDNA) are present in the peripheral blood of cancer patients (see **Figure 1**).



	biopsy	CTC	ctDNA
invasive	+	–	–
all patients eligible	–	+	+
instrumentation required	+	+	–
WGA required	–	+	+/–
RNA profiling	+	+	–
research applicability	+++	++	+
biomarker applicability	–	++	+++

Figure 1 Comparison of liquid biopsies versus needle biopsies for investigating cancer. CTC, circulating tumor cells; ctDNA, circulating (cell-free) tumor DNA; WGA, whole genome amplification. (Reproduced from Wyatt AW & Gleave ME [2015] *EMBO Mol Med* 7:878–894; PMID 25896606. © 2015 The Authors. Published under the terms of the CC BY 4.0 license.)

Retrieving and analyzing tumor DNA from peripheral blood is not straightforward because of the limited quantity of tumor material in the circulation (1 ml of blood has 10 million leukocytes but maybe only 1 CTC; and the fraction of circulating cell-free DNA that originates from tumors may often be very low). Currently, specialized techniques such as droplet digital PCR are required to amplify the DNA prior to sequence analysis.

Of the numerous applications of liquid biopsies in cancer, a very important one will be in monitoring the evolution of treatment resistance to detect, for example, secondary mutations that arise in tumors to prevent a treatment drug working. By identifying the drug-resistance mutation before clinical signs of disease progression develop, it may be possible to quickly make a compensatory adjustment to the treatment.

The promise of combinatorial drug therapies

Because of tumor heterogeneity, is targeted drug therapy always doomed to eventual failure? One approach that has great potential is combinatorial therapy, that is, using combinations of treatments that act differently. It conceivably might even have the potential to lead to actual cures for some cancer patients, rather than simply temporary remissions. A successful template has already been provided by the recent success in treating human immunodeficiency virus (HIV). Just like tumor cells, the HIV virus is highly mutable and can quickly mutate to resist any

individual antiviral drug. But the highly active antiretroviral therapy (HAART) strategy used a combination of different antiretroviral drugs. Because the chances that individual viruses could mutate to become simultaneously resistant to two or three drugs are low, HAART has been much more successful than previous HIV treatments.

Future precision oncology might involve the simultaneous use of multiple agents and drugs that target diverse vulnerabilities of cancer cells before resistance has a chance to develop. There are potentially huge numbers of possible permutations. One might target upstream and downstream components in a pathway known to be important in the development of specific cancers, or components in parallel pathways. There are different types of treatment, including standard drug therapies and immunotherapies. By 2020 more than 5000 clinical trials were ongoing globally to assess the clinical benefits from new combination therapies. Because the possibilities to combine treatments dramatically outnumber the patients available to enroll in clinical trials there is an urgent clinical need for rational cancer treatment combinations.

SUMMARY

- Cancers are diseases in which there is an unregulated increase in cell growth that leads to cells invading neighboring tissues and spreading to distant sites in the body.
- The genetic contribution to cancers predominantly occurs through somatic mutations. Germline mutations may result in inherited cancers, but even in these cases additional somatic mutations are required for cancers to form.
- Cancer development occurs only after a series of succes sive regulatory controls have gone wrong in cells, leading to increased cell proliferation or reduced apoptosis.
- Cancers are primarily diseases of later life, because it takes time for multiple cell controls to be disrupted.
- Tumors originate ultimately from a single cell but they are genetically heterogeneous. Descendants of the founder cell can acquire genetic mutations that afford a growth advantage; they form a

dominant subclone that is then surpassed in growth by successive subclones (which acquire additional mutations conferring further growth advantages).

- Cancers develop by accelerating mutation in two dif ferent ways: by conferring a growth advantage in cells, and by destabilizing the genome to increase the probability of later mutations.
- In some types of cancer, undifferentiated cells are found with stem cell properties; they can self-replicate and also give rise to more differentiated cells within the tumor. Genetically different cancer stem cells may also arise by clonal evolution.
- Tumors most likely originate from cells that already have a high proliferative capacity, such as stem cells or rapidly multiplying and poorly differentiated embryonic tissues. But genetic and epigenetic changes in differentiated cells may also cause these cells to become progressively more plastic (less differentiated) and progressively acquire other characteristics of cancer cells.
- Intratumor heterogeneity includes not just genetically different descendants from a single founding cell, but also non-tumor cells that are recruited to the tumor microenvironment, including some types of infiltrating immune cell.
- Cancer is a contest between Darwinian natural selec tion operating at the level of the individual (over generations) and at the level of the cell (within a single individual). Although cancer cells can successfully proliferate and form tumors within a person, they cannot leave progeny beyond the life of their host; tumorigenesis processes must start afresh in a new individual.
- Cancer cells usually contain thousands of somatic mutations. A small number, often from one to five or six, are driver mutations that are crucially important in cancer development and are positively selected. The rest are chance (passenger) mutations resulting from genomic instability.
- Cancer genes can be grouped into two classes accord ing to how they work in cells. In some cases, mutation of a single allele is sufficient to make a major contribution to the development of cancer. For other

cancer genes, both alleles need to be inactivated to make a significant contribution to cancer.

- Oncogenes are dominantly acting cancer genes that arise through some activating mutation in one allele of a normal cellular "proto-oncogene". Classical protooncogenes typically work in growth signaling pathways to promote cell proliferation or inhibit apoptosis.
- Proto-oncogenes can be activated to become onco-genes by acquiring gain-of-function mutations; by being over-expressed as a result of gene amplification; or through activated expression resulting from a trans-location (which repositions a transcriptionally silenced gene so that it comes under the control of transcription-activating regulatory elements).
- Classical tumor suppressor genes are recessively act ing cancer genes in which the inactivation of both alleles promotes cell proliferation or inhibits apoptosis. Additional tumor suppressors work in other areas such as in genome maintenance.
- The two-hit hypothesis describes how cancer devel ops from two successive inactivating mutations in a tumor suppressor gene. It explains why dominantly inherited cancers are recessive at the cellular level (the first mutation occurs in the germ line and so there is a very high chance that the second allele is inactivated in at least one cell in the body to form a tumor). In sporadic cancers of the same type, both the first and second inactivating mutations occur in a somatic cell.
- Genome instability ensures additional mutations for natural selection to work on to drive tumor formation. It often manifests as chromosomal instability (resulting in aneuploidies, translocations, and so on) but can also be apparent at the DNA level as microsatellite instability (resulting from mutations in genes that work in mismatch DNA repair).
- Epigenetic dysregulation is important in both cancer initiation and cancer progression. It can be induced by genetic changes (notably mutation in genes that make epigenetic regulators) or by tissue inflammation causing altered cell signaling that results in altered chromatin states.

- Aberrant chromatin states produced by epigenetic dysregulation can allow cancer cells to become unspecialized (poorly differentiated) and can silence alleles of cancer-susceptibility genes. Additionally, DNA hypomethylation can result in widespread chromo-some instability.
- Genome-wide gene expression profiling of tumors can subdivide cancers of the same type, such as breast carcinomas, into different groups with different biological characteristics and different drug responses.
- Two tumors of the same type show very different mutational spectra —the great majority of passenger mutations are often distributed randomly across the genome; although some key cancer genes might be mutated in both tumors, other driver mutations may be located in different cancer-susceptibility genes.
- Tumors evolve, so cells in different regions of the same tumor can show regional mutational differences; metastatic cells typically share mutations that distinguish them from the primary tumor.
- Human cancer-susceptibility genes have been identi fied by analyzing associated chromosome breakpoints or associated changes in copy number (oncogene amplification, or loss of heterozygosity in the case of tumor suppressor genes); by studying candidate genes suggested by analyses of experimental organisms; and by exome or genome sequencing.
- In targeted cancer therapies, a drug or other treatment agent is directed at counteracting the effects of a specific genetic mutation that is known to be crucial for development of the cancer.
- Recurrence of tumors may be driven by cancer stem cells that are comparatively resistant to therapy.
- After initial success in shrinking tumors, cancer therapies often fail, causing a clinical relapse. Tumor cells evolve to become resistant to the drug as a result of natural selection (which promotes the growth of tumor cells that develop mutations to combat the effects of the drug).
- Tumors often develop drug resistance by changing the conformation of the drug target so that the drug is sterically hindered from binding

to it, by amplifying the gene encoding the drug target or by activating an alternative pathway that bypasses the effect on the drug target.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

Cancer biology

Weinberg, RA (2014) *The Biology of Cancer*, 2nd ed., Garland Science.

General molecular characteristics of cancer

Hanahan D & Weinberg R (2011) Hallmarks of cancer: the next generation. *Cell* 144:646–674; PMID 21376230.

Shay JW & Wright WE (2011) Role of telomeres and telomerase in cancer. *Semin Cancer Biol* 21:349–353; PMID 22015685.

Trybek T (2020) Telomeres and telomerase in oncogenesis. *Oncol Lett* 20:1015–1027; PMID 32724340.

Vaupel P, Multhoff G (2021) Revisiting the Warburg effect: historical dogma versus current understanding. *J Physiol* 599:1745–1757; PMID 33347611.

Cancer evolution, cancer stem cells, and intratumor heterogeneity

Burrell RA (2013) The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 501:338–345; PMID 24048066.

Clevers H (2011) The cancer stem cell: premises, promises, and challenges. *Nature Med* 17:313–319; PMID 21386835. [For a follow-up in 2017 see also PMID 28985214.]

Gerlinger M (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366(10):883–892; PMID 22397650.

Greaves M & Maley CC (2012) Clonal evolution in cancer. *Nature* 481:306–313; PMID 22258609.

Hausser J & Alon U (2020) Tumor heterogeneity and the evolutionary trade-offs of in cancer. *Nat Rev Cancer* 20:247–257; PMID 32094544.

Magee JA (2012) Cancer stem cells: impact, heterogeneity and uncertainty. *Cancer Cell* 21: 283–296; PMID 22439924.

Pon JR & Marra MA (2015) Driver and passenger mutations in cancer. *Annu Rev Pathol Mech Dis* 10: 25–50; PMID 25340638.

Reiter JG (2019) An analysis of genetic heterogeneity in untreated cancers. *Nat Rev Cancer* 19:639–650; PMID 31455892.

Trumpp A, Haas S (2014) Cancer stem *cell*s: the adventurous journey from hematopoietic to leukemic stem cells. *Cell* 185:1266–1270; PMID 35385684.

Oncogene activation and chromosome aberrations in cancer

Hnisz D (2016) Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* 351:1454–1458; PMID 26940867.

Mertens F (2015) The emerging complexity of gene fusions in cancer. *Nat Rev Cancer* 15:371–381; PMID 25998716.

Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer. https://mitelmandatabase.isb-cgc.org/

Roukos V (2013) The cellular etiology of chromosome translocations. *Curr Opin Cell Biol* 25:357–364; PMID 23498663.

Storlazzi CT (2010) Gene amplification as double minutes or homogenously staining regions in solid tumors: origin and structure. *Genome Res* 20:1198–1208; PMID 20631050.

Tumor suppressor genes

Berger AH (2011) A continuum model for tumor suppression. *Nature* 476:163–169; PMID 21833082.

Berger AH & Pandolfi PP (2011) Haplo-insufficiency: a driving force in cancer. *J Pathol* 223:137–146; PMID 21125671.

Freed-Pastor WA & Prives C (2012) Mutant p53: one name, many proteins. *Genes Dev* 26:1268–1286; PMID 22713868.

Knudson AG (2001) Two genetic hits (more or less) to cancer. *Nat Rev Cancer* 1:157–162; PMID 11905807. [A historical perspective of the development of the twohit tumor suppressor hypothesis.]

Sebastian C (2012) The histone deacetylase SIRT6 is a tumor suppressor that controls cancer metabolism. *Cell* 151(6):1185–1199.

Solimini NL (2012) Recurring hemizygous deletions in cancer may optimize proliferative potential. *Science* 337:104–109; PMID 22628553.

Genome instability and epigenetic dysregulation in cancer

Darwiche N (2020) Epigenetic mechanisms and the hallmarks of cancer: an intimate affair. *Am J Cancer Res* 10:1954–1978; PMID 32774995.

Feinberg AP (2016) Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nature Rev Cancer* 17:284–299; PMID 26972587.

Flavahan WA (2017) Epigenetic plasticity and the hallmarks of cancer. *Science* 357:eaal2380; PMID 28729483.

Pena-Diaz J & Jiricny J (2012) Mammalian mismatch repair: error-free or error-prone? *Trends Biochem Sci* 37:206–214; PMID 22475811.

Roy R (2012) BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat Rev Cancer* 12:68–78; PMID 22193408.

Schwitalla S (2013) Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem *cell*-like properties. *Cell* 152: 25–38; PMID 23273993.

Shen H & Laird PW (2013) Interplay between the cancer genome and epigenome. *Cell* 153:38–55; PMID 23540689.

Cancer genomics and transcriptomics

Cieslik M & Chinniyan AM (2020) Global cancer genomics project comes to fruition. *Nature* 578:39–40. [Provides a commentary on six key papers

published in the same issue from the Pan Cancer Analysis of Whole Genomes consortium.]

Garraway LA & Lander ES (2013) Lessons from the cancer genome. *Cell* 153:17–37; PMID 23540688.

Hong M (2020) RNA sequencing—new technologies and applications in cancer research. *J Hematol Oncol* 13:166; PMID 33276803.

Kim N (2020) Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nature Commun* 11:2285; PMID 32385277.

Kim B (2020) Advancing cancer research and medicine with single-cell genomics. *Cancer Cell* 37:456–470; PMID 32289270.

Martinez-Jimenez F (2020) A compendium of mutational cancer driver genes. *Nat Rev Cancer* 20:555–572; PMID 32778778.

Parsons J & Francavilla C (2020) Omics approaches to explore the breast cancer landscape. *Front Cell Dev Biol* 7:395; PMID 32039208.

Paull EO (2021) A modular master regulator landscape controls cancer transcriptional identity. *Cell* 184:334–351; PMID 33434495.

Vogelstein B (2013) Cancer genome landscapes. *Science* 339:1546–1558; PMID 23539594.

Cancer therapeutics and monitoring strategies

Al-Lazikani B (2012) Combinatorial drug therapy for cancer in the post-genomic era. *Nature Biotechnol* 30:1–13; PMID 22781697.

Crowley (2013) Liquid biopsy: monitoring cancer-genetics in the blood. *Nat Rev Clin Oncol* 10:472–484; PMID 23836314.

Dancey JE (2012) The genetic basis for cancer treatment decisions. *Cell* 148:409–420; PMID 22304912.

June CH (2018) CAR T cell immunotherapy for human cancer. *Science* 369:1361–1365; PMID 29567707.

McDermott U (2011) Genomics and the continuum of cancer care. *N Engl J Med* 364:340–350; PMID 21268726.

*Nature Med*icine Focus on targeted cancer therapies (2013) Nature Med 19:1380–1464. [A collection of various reviews in this area in the November 2013 issue.]

Ward RA (2020) Challenges and opportunities in cancer drug resistance. *Chem Rev* 6:3297–3351; https://dx.doi.org/10.1021/acs.chemrev.0c00383

11

Genetic and genomic testing in healthcare

Practical and ethical aspects

DOI: 10.1201/9781003044406-11

CONTENTS

We end this book by considering how the ever-expanding knowledge of our genome, our genes, and genetics is being used in genetic testing to make a

positive impact on the health of society. Our ability to scrutinize and interpret genetic variation in health and disease has certainly expanded substantially since the last edition of this book. And there have been new developments in treating disease. We discuss how these advances raise ethical considerations for practice, and how further developments in both diagnosis and treatment may raise different versions of the ethical issues.

In some previous chapters we looked at how genetics and genomics are illuminating our understanding of the molecular basis of disease, and how this knowledge has brought about significant—sometimes profound—changes in how we diagnose and treat human disease. For many genetic disorders there remains no adequate treatment; in general, genetic approaches to treatment have advanced much more slowly than our ability to diagnose a genetic cause. Innovative treatments are emerging, however, and it is likely that there will be many more changes over the coming decade.

Genetic testing has been used in the clinic since the late 1960s. The subsequent DNA cloning revolution allowed rapid developments in DNA-based diagnosis. Initially used in just a few medical settings, notably clinical genetics, DNA technologies were then democratized by PCR, an inexpensive DNA technology that was very easy to use and was rapidly taken up. As well as being extensively used in clinical genetics services, PCR-based testing became the standard way of identifying pathogens and so is a key tool used by microbiologists and virologists. And it has come to be increasingly used in other medical specialties, such as hematology and oncology. Of course, there has also been a revolution in DNA sequencing.

As the technologies become exponentially faster and cheaper the entire clinical approach is changing. First, genetic testing is now available, and useful, to all the major divisions of medicine—it is certainly no longer the preserve of specialties labeled “clinical genetics”. Secondly, previous strategies of using a phenotype to determine which bits of a genotype to assess have become inverted: whole genome assays are quite often the first step from which predictions about phenotype might now be made.

Given its speed and affordability, whole genome genetic testing is being offered by more and more private companies direct to the consumer (DTC). DTC genetic testing comes with or without health interpretations, and accompanying healthcare professional input may be totally lacking or minimal. Given that the popular discourse around genetic information is often strikingly optimistic—"your DNA is your blueprint" and so on—there is potential for underestimating the complexity of a DTC output in terms of predictions about future health.

In Section 11.1 we give an overview of genetic testing before describing the technology of genetic testing for detecting chromosome abnormalities and large to moderate-scale DNA copy number variation (Section 11.2), and testing for point mutations and DNA methylation changes (Section 11.3). In Section 11.4, we describe how genetic services are organized and the practical applications of genetic testing. (Note that we have previously described applications in pharmacogenetic testing within the context of treatment for genetic disorders; readers interested in this application should consult Section 9.2.)

Finally, in Section 11.5, we consider the range of ethical questions and impacts on society that might arise through the practices of genetic testing and certain applications of genetic technologies to treat disease. We then go on to offer some thoughts and possible directions to turn to so that such issues can be resolved or ameliorated in practice.

11.1 AN OVERVIEW OF GENETIC TESTING

Although the title of this chapter refers to genetic testing in healthcare, it is important to consider that genetic testing requested for other reasons may intersect with healthcare analyses. Genetic testing for crime scene analyses, identity checking, or determining whether biological relationships have been misattributed, might each be conducted for legal reasons, but they may also provide important information for healthcare. Companies offering direct-to-consumer (DTC) genetic testing increasingly sell ancestry tests to pique consumer interest with certificates of ancestral make-up, but often

such data are also analyzed to make healthcare predictions, and so may intersect with healthcare questions. Genetic testing is also important in understanding normal genetic variation in different human populations. And, increasingly, genetic testing is being used to analyze the genetic contribution to common, complex genetic disease rather than being limited to the rare disease diagnoses it was focused on just a few decades ago.

The genetic testing outlined in this chapter is primarily concerned with detecting the relatively small portion of human genetic variation that confers susceptibility to disease. There are different general strategies for carrying out the testing, and different levels and environments at which it is carried out.

The different source materials and different levels of genetic testing

The source material for genetic testing can be cells (usually blood, tumor, skin, embryonic, or fetal cells; see **Table 11.1**) or body fluids (blood, urine), stools, and even exhaled breath (increasingly used in assaying certain cancer biomarkers). In addition, testing is sometimes carried out on archived material (often stored blood or tumor samples) from deceased persons to provide information that can be of clinical help to surviving family members.

TABLE 11.1 SOURCES OF MATERIAL FOR GENETIC TESTING

Source of cells/DNA/RNA	**Type of testing or scereening**
EMBRYONIC/FETAL	
Single cell from a blastomere or a few cells from a blastocyst	preimplantation diagnosis
Fetal DNA in maternal blood	prenatal diagnosis, as early as 6 weeks (testing for paternal alleles). Fetal sexing

Source of cells/DNA/RNA	**Type of testing or scereening**
Chorionicvillus	prenatal diagnosis at about 9-14 weeks
Amniotic fluid	prenatal diagnosis at about 15-20 weeks
Umbilical cord blood	prenatal diagnosis at about 18-24 weeks
ADULT/POSTNATAL	
Peripheral blood	screening for heterozygote carriers.Testing for defined heterozygous carrier genotype. Pre-symptomatic genotype screening or testing. Identity testing (DNA profiling). Testing for chromosome abnormalities
Mouthwash/buccal scrape	
Biopsy of skin/muscle/other tissue	RNA-based testing
Tumor biopsy	cancer-associated genotypes or gene expression patterns
Guthrie card	neonatal screening
ARCHIVED MATERIAL FROM DECEASED PERSONS	
Pathological specimens	genotyping
Guthrie card	possible source of DNA from a deceased individual (not all of the blood spots on the card might have been used up in neonatal screening)

Like other clinical tests, genetic tests may be conducted on individuals, couples, families, communities, or whole populations. They may be conducted at different levels for different purposes, whether informing about genetic risk at the prenatal diagnosis level, at preimplantation, at pre-conception (to inform a decision about future reproduction), at the level of managing an existing genetic condition, and so on. And the tests may be of different types, according to whether the object is to detect a *specific* pre-

defined genetic abnormality for some purpose, or to *scan* for a variety of *possible* genetic changes, that then must be further investigated to see if any of them represent a convincing pathogenic change. See **Figure 11.1** for a visual representation with some illustrative examples of the methods used.

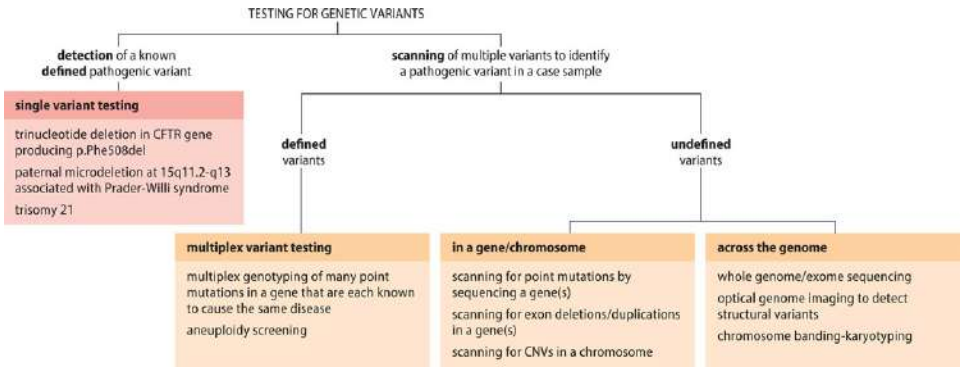


Figure 11.1 Some examples to illustrate how genetic testing to is used to confirm or define pathogenic DNA variants. Detection of a single pathogenic variant may be conducted to confirm a previously defined genetic variant or abnormality present in another family member or suggested by the phenotype, or in management of a cancer. Alternatively, genetic testing is conducted to define an unknown causative mutation. If the disease gene locus and many previously identified causative mutations are already known, multiplex genotyping of several, or many, *known* causative mutations may be carried out to scan for the pathogenic variant in an affected individual. Other types of scanning for a causative mutation may be much broader because the pathogenic variants are undefined; here, the object is to define the causative variant. In cases where there is uncertainty over the disease locus, genome-wide scans can be carried out; they may identify many candidate pathogenic mutations that first need to be evaluated. CNVs, copy number variants of large sequences.

Genetic testing started in the 1960s with looking for chromosome abnormalities by examining stained chromosome preparations under the microscope. Linkage analysis and mutation testing for rare highly penetrant mutations within genes underlying monogenic diseases followed in the 1980s with more and more genes being identified, and thus testable, over the next few decades. These services were focused on advising others—

existing family members or possible future ones via pregnancy investigations—once a disease or phenotype had come to light.

As techniques improved, the possibility of genetic testing was extended from the prenatal setting (from cells taken at different stages of pregnancy), to also include preimplantation genetic diagnosis (PGD). The latter is conducted in the context of *in vitro* fertilization: tests are carried out to decide which embryos lack the genetic abnormality being investigated and can thus be implanted. More recently, preconceptual testing for a panel of individually rare autosomal recessive conditions offers improved reproductive choices before any fertilization takes place.

Predictive genetic testing and genetic screening

Predictive genetic testing is increasingly offered to family members as more and more genetic diagnoses are made. Often, such tests are directed at members of a family with a history of a late-onset single-gene disorder, and the aim is to predict whether an asymptomatic member of the family is at risk of the disorder at some *future* point. A positive test result may offer the opportunity to take some medication and/or alter lifestyle factors in the meantime to reduce the disease risk; a negative test result provides reassurance that the predisposing genetic factor has not been inherited.

Genetic testing is also carried out on apparently asymptomatic individuals in communities and populations to identify individuals carrying harmful variants (*genetic screening*). The aim is usually to identify a high-risk subset of the population who can then be offered additional specific testing (such as follow-up prenatal diagnosis after identifying couples who are both carriers of a recessive condition). Unless the diseases are dominated by a few known types of genetic variant, we typically do not know what variants may be present in the individuals. As a result, genetic screening quite often involves assays of gene products or biomarkers associated with the pathogenesis, such as altered metabolites.

Direct versus indirect genetic testing

As previously summarized in Figure 7.1, genetic disorders arise through pathogenic DNA sequence changes or copy number changes. Direct genetic testing means that the test assays for the presence of a causative sequence or copy number mutation. For a very few disorders, exceptional mutational homogeneity allows us to predict the genotype, such as for sickle-cell disease and Huntington disease. But for the vast majority of single-gene disorders, if we do not have prior knowledge of the molecular pathology in a family, we first need to screen candidate genes to identify the causative mutation. After that, we can carry out a direct assay to determine the risks for relatives by seeing whether they carry the disease-associated variant or not.

To identify the causative mutation some type of *gene analysis* is carried out. Sometimes, multiple genes are analyzed, as for certain single-gene disorders (the Lynch syndrome case study profiled in Clinical Box 13 on page 394 gives an example), and for some common cancers. In response to a person enquiring about their family history of breast and ovarian cancer, for example, DNA from an affected relative can be assayed for point mutations/copy number variation in a panel of genes known to cause such familial disease. Candidate pathogenic DNA variants can then be investigated using various approaches described below to assess pathogenicity.

In *indirect genetic testing* the assay does not screen for the pathogenic variant directly; instead, it assays some other factor that is genetically linked to the variant or is a direct consequence of the variant. At the end of the day, a test is sufficient if it lets us know whether or not a particular gene variant is present even if we do not directly assay that variant.

In the past, indirect tests were often carried out using linkage analysis. Polymorphic DNA markers that mapped very close to a disease gene locus would be assayed in order to *infer* the inheritance of disease alleles using the same approach as previously illustrated in Figure 8.2. Such tests had small error rates because of the chance of recombination between marker

and disease locus. With the sequencing of the human genome and the development of rapid gene sequencing techniques, indirect linkage analyses have almost become obsolete.

Some indirect genetic testing assays some consequence of genetic variation, rather than the genetic variation itself. The testing might seek evidence of a gene product, or a characteristic disease-associated biomarker, such as an abnormally elevated metabolite. Sometimes a functional assay can be used. For recessive disorders, a single functional assay might be sufficient to detect a loss of function and can be conveniently carried out in cultured cells in which the gene is expressed. See **Table 11.2** for some examples of genetic testing that assay the consequence of genetic variation.

TABLE 11.2 EXAMPLES OF INDIRECT GENETIC TESTS THAT ASSAY SOME CONSEQUENCE OF GENETIC VARIATION

Assay level	**Example**
Gene product	Proteins: may be assayed by immunohistochemistry (as in Lynch syndrome in Clinical Box 13 on page 394). RNA:as in identifying translocations by targeted sequencing of transcripts of fusion genes
Genome instability	Microsatellite instability in Lynch syndrome. Because of a deficiency in mismatch repair, microsatellites across the genome show aberrant profiles.

Assay level	Example
Functional assay	Tests for various genes encoding enzymes may be assayed by enzyme assays. DNA testing for the DNA repair disorder Fanconia anemia is often conducted by an *in vitro* DNA repair assay (DNA detection is complicated because the disease can be caused by a mutation in any one of 15 different genes, some with large numbers of exons). Cultured lymphocytes from an individual are treated with a DNA interstrand cross-linking agent (often diepoxybutane or mitomycin C), and examined to identify chromosomal aberrations resulting from defective repair of the induced DNA cross-linking.

How genetic tests can be evaluated

Given the increasing availability in healthcare practice of genetic tests designed to identify a specific class of DNA change, standards for the assessment of their performance are important. Several frameworks for such evaluation have been proposed. The ACCE model, is one most referenced in the literature, proposes four main criteria:

- **A**nalytical validity: how well does the test assay measure what it claims to measure?
- **C**linical validity: how well does the test predict the projected health outcome?
- **C**linical utility: how useful is the test result?
- **E**thical validity: how well does the test meet the expected ethical standards?

The evaluation, particularly for predictive tests or tests for not very penetrant genes, can be more complex than implied by the four categories above. Rapid development and marketing may mean that there is not yet

sufficient evidence to answer all these questions, and lack of public health/population level data may mean that ascertainment of particular test results selects for additional familial factors not measured by the test. Performing high-quality randomized controlled trials to demonstrate utility is often difficult, and the lack of evidence of effectiveness of a test may affect an evaluation of cost-effectiveness.

The ACCE process has been used by policy makers to decide about genetic testing for particular disorders using a standard set of 44 targeted questions (https://www.cdc.gov/genomics/gtesting/acce/index.htm) that address disorder, testing, and clinical scenarios, as well as analytic and clinical validity, clinical utility, and associated ethical, legal, and social issues.

The analytical validity of the test is determined by two key performance indicators, as follows:

- the **sensitivity**, the proportion of all people with the condition who are correctly identified as such by the test assay
- the **specificity**, the proportion of all people who do not have the condition and who are correctly identified as such by the test assay.

See **Table 11.3** for a worked example and for how related measures are defined.

TABLE 11.3 PARAMETERS RELATING TO THE ANALYTICAL VALIDITY OF A TEST

		CONDITION	
		Present	Absent
TEST	+ve	a (90)	c (30)
	-ve	b (10)	d (1870)

Sensitivity	a/*a*+*b*	(90/100 = 90 %)
Specificity	d/*c*+*d*	(1870/1900 = 98.4 %)
False positive rate	c/*a*+*c*	(30/120 = 25 %)
Positive predictive value	a/*a*+*c*	(90/120 = 75 %)
False negative rate	b/*b*+*d*	(10/1880 = 0.5 %)
Negative predictive value	d/*b*+*d*	(1870/1880 = 99.5 %)

Numbers in parentheses are specific values for illustrative purposes only, drawn from 100 people with a hypothetical condition and 1900 lacking the condition. The false positive rate is the proportion of people who test positive for the factor being assayed but do not have the condition. The false negative rate is the proportion of people who test negative for the factor being assayed but who have the condition. The positive predictive value is the proportion of people testing positive who have the condition. The negative predictive value is the proportion of people testing negative who do not have the condition. Note that the sum of the false positive rate and the positive predictive value is always 100 %, as is the sum of the false negative rate and the negative predictive value.

The ACCE process produces important, but under-evaluated, by-products. First, it identifies where the gaps in knowledge are in the natural history of a disease (which is important for future research agendas). Secondly, it can identify where the implementation gaps are. After the test findings are given, downstream recommendations can be made, such as screening or interventions. But of course we need to know about, and then reduce, the barriers to implementing these downstream recommendations, as well as effects of the recommended screening or intervention.

11.2 GENETIC TESTING FOR CHROMOSOME ABNORMALITIES AND PATHOGENIC STRUCTURAL VARIATION

In Section 7.4 we detailed the two fundamental classes of chromosome abnormality (large-scale DNA changes that can be detected by standard karyotyping using chromosome banding techniques). They are: numerical abnormalities (in which abnormal chromosome segregation leads to aneuploidy, with fewer or more chromosomes copies than normal); and structural abnormalities (in which standard karyotyping by chromosome banding reveals chromosome rearrangements that produce large-scale deletions, duplications, inversions, or translocations).

Of course, disease can also be caused by structural abnormalities below the limit of detection of standard chromosome banding techniques. They mostly manifest as copy number variants (CNVs), comprising deletions and duplications from over 50 bp to a few Mb of DNA. They usually cause disease by eliminating one or more genes or by inactivating a gene as result of deletion or duplication or one or more exons or gene control regions. Various molecular genetic techniques can be used to screen/detect such DNA changes.

TABLE 11.4 AN OVERVIEW OF (A) MAJOR TECHNIQUES USED TO SCREEN FOR, OR CONFIRM, CHROMOSOME ABNORMALITIES AND PATHOGENIC LARGE COPY NUMBER VARIANTS, AND (B) ONLINE RESOURCES TO ASSIST IN THEIR INTERPRETATION

(A)	TECHNIQUE	APPLICATION
	standard karyotyping (chromosome banding)	General method for screening for chromosome abnormalities. Detailed in **Box 7.2** on pp. 204–5. Often now used as a back-up method.
	quantitative fluorescence PCR (QF-PCR)	The front-line method to screen for the common aneuploidies in prenatal diagnosis.
	chromosome SNP microarray analysis	The most commonly used type of chromosome microarray analysis. The method of choice for screening for large deletions and duplications across the genome. Also, a confirmatory method for screening for aneuploidies in prenatal diagnosis.

* The older alternative of Southern blot-hybridization is virtually obsolete now, but still used in some labs to detect large deletions in facioscapulohumeral dystrophy – for an example, see Figure 2 in Clinical Box 3 on p. 171. It, and triplet repeat-primed PCR, can be used to detect very large expansion of repeats n disorders such as Fragile X syndrome and myotonic dystrophy.

chromosome FISH (fluorescence in situ hybridization)	Often used to confirm regions of chromosome deletions/duplications identified by chromosome microarray analysis. Also used to screen for amplification of oncogenes (Figure 10.7A on page 377) and in detecting translocations, especially common oncogenic translocations.
RNA fusion screening	RNA fusion panels permit general screening for cancer-causing translocations (using targeted RNA sequencing to identify transcripts of many possible oncogenic fusion genes arising via translocation).
multiplex ligation-dependent probe amplification (MLPA)*	Especially used to test for gene variants where one or more exons are duplicated or deleted. Commercial kits are available for many genes, but this is not an easy method to carry out where kits are not available.
droplet-digital PCR (ddPCR)	A type of quantitative PCR. Offers highly sensitive accurate quantitation of CNVs. Highly versatile method. Needs dedicated PCR machine.
genome-wide sequencing	A universal screen that can identify structural variants across the genome (as well as point mutations).

* The older alternative of Southern blot-hybridization is virtually obsolete now, but still used in some labs to detect large deletions in facioscapulohumeral dystrophy – for an example, see Figure 2 in Clinical Box 3 on p. 171. It, and triplet repeat-primed PCR, can be used to detect very large expansion of repeats n disorders such as Fragile X syndrome and myotonic dystrophy.

	optical genome mapping	A universal screen for structural variation.This new and different approach can be used to detect large and small structural variants, copy number variations, and complex rearrangements.
(B)	ONLINE RESOURCE TO ASSIST INTERPRETATION	DESCRIPTION
	Decipher database	At http://decipher.sanger.ac.uk. The Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources collects clinical information about chromosomal microdeletions, microduplications, insertions, translocations, and inversions, and displays this information on the human genome map.
	ClinVar	At https://www.ncbi.nlm.nih.gov/clinvar/intro/. Maps relationships between human variations and phenotypes.
	dbVar	At https://www.ncbi.nlm.nih.gov/dbvar/. NCBI's database of human genomic structural variation. Documents large variants (>50 bp), including: insertions, deletions, duplications, mobile elements, translocations, and complex variants.

* The older alternative of Southern blot-hybridization is virtually obsolete now, but still used in some labs to detect large deletions in facioscapulohumeral dystrophy – for an example, see Figure 2 in Clinical Box 3 on p. 171. It, and triplet repeat-primed PCR, can be used to detect very large expansion of repeats n disorders such as Fragile X syndrome and myotonic dystrophy.

In **Table 11.4** **Part (A)** we give an overview of the principal techniques used to screen for chromosome abnormalities and large-to moderate-scale

CNVs and their main applications. The interpretation of these tests are aided by various databases and online resources (**Table 11.4 Part (B)**).

Screening for fetal aneuploidies using quantitative fluorescence PCR

The commonest aneuploidies are trisomies 13, 18, and 21 and various types of abnormal sex chromosome number, and prenatal testing has been available for some time. Initial techniques required chromosome culturing so that results took a minimum of two weeks to grow the appropriate cells. More recently, quantitative fluorescence PCR has become the screening method of choice: it offers a much more rapid turnaround, often 24 hours, either by detecting abnormalities on fetal scanning, or increasingly through non-invasive prenatal testing (NIPT) offered to the population of pregnant women, as described below.

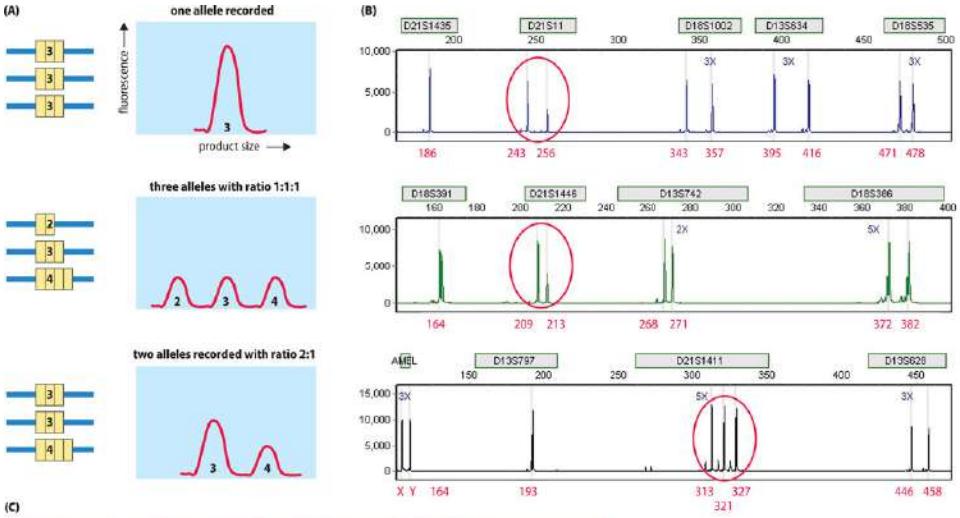
Quantitative fluorescence PCR (QF-PCR) is fast, robust, highly accurate, and largely automated. Several pairs of fluorescently labeled primers are used in a *multiplex PCR*—the idea is to simultaneously amplify multiple polymorphic markers on the chromosomes most frequently involved in aneuploidies. For each marker, the amplification products will fall within a characteristic size range of different lengths; as required, two or more markers that have overlapping allele sizes can be distinguished by labeling them using fluorophores that fluoresce at different wavelengths.

Certain polymorphic short tandem repeat polymorphisms are usually selected, often based on tetranucleotide or pentanucleotide repeats to maximize the length difference between alleles. Fluorescently labeled products from the *exponential phase* of the PCR reaction (Figure 3.4) are separated according to size by electrophoresis through long and extremely thin tubes containing polyacrylamide (capillary electrophoresis). That happens in a commercial DNA analyzer of the type used in capillary DNA sequencing: a detector at a fixed position records the intensity of fluorescence signals as fragments migrate through the capillary tubes and

past the detector (Box 3.3 on page 73 describes the principle of capillary electrophoresis).

Autosomal aneuploidies

To monitor the common autosomal aneuploidies, the QF-PCR screen uses highly polymorphic short tandem repeat markers. An individual marker might not always be informative: in a trisomy, for example, the marker might show identical repeat numbers for all three chromosome copies, just by chance, resulting in an uninformative, single PCR product. The most informative situation occurs when the marker exhibits different numbers of repeats on the three chromosomes. But quite often only two length variants are recorded for a single marker; then quantification becomes important (**Figure 11.2A**). Because four or more different markers are used per chromosome, however, there is little difficulty with interpretation (two or more markers are often informative for each chromosome—see Figure 11.2B,C for a practical example).



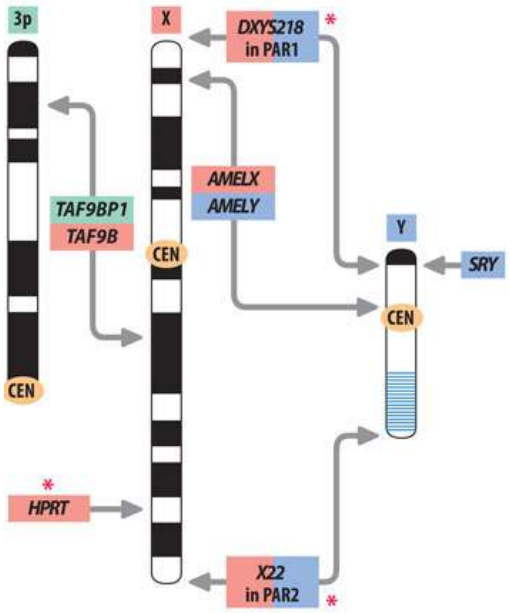
(C)

Marker	Alleles	Allele Length	Peak Area
AMEL	2	X:Y	15263:15151
D13S628	2	446:458	16705:14667
D13S634	2	395:416	11761:11821
D13S742	2	268:271	18747:16855
D13S797	1	193	55148
D18S1002	2	343:357	10439:9373
D18S386	2	372:382	8719:8742
D18S391	1	164	31165
D18S535	2	471:478	12260:12054
D21S11	3(2:1)	243:256	28511:12207
D21S1411	3(1:1:1)	313:321:327	12058:11870:11820
D21S1435	1	186	34476
D21S1446	3(2:1)	209:213	37508:18331

Figure 11.2 Autosomal trisomy screening using QF-PCR. (A) Interpreting marker data (on the right) from an imagined locus with three common alleles that have two, three, or four tandem repeats, as illustrated on the left. The top trace is uninformative (just one length variant is recorded). The middle trace is fully informative: the presence of three alleles with different lengths strongly suggests trisomy. The bottom trace is suggestive of trisomy: two length variants are evident, but the fluorescence associated with allele 3 seems to be approximately twice that associated with allele 4 (the area under the peaks is normally used for quantitation). (B) A practical example. The output shows traces for three sets of markers (shown in blue at top, green in the middle, and black at the bottom) that collectively represent assays for microsatellite markers on chromosomes 13, 18, and 21 (the three autosomes associated with viable trisomies), plus control X and Y markers from the amelogenin genes (see Figure 11.3). The data highlighted by red ovals strongly suggest trisomy in this individual: three alleles of different sizes for *D21S1411,* and a 2:1 ratio for the two length variants for each of *D21S11* and *D21S1446*. The other chromosome 21 marker, *D21S1435,* is uninformative, showing only one length variant, which is presumably due to three alleles of identical lengths. (C) The calculation of peak areas and interpretation by

SoftGenetics software (rows highlighted in gray are significant). (B,C, data courtesy of Jerome Evans, NHS Northern Genetics Service, Newcastle upon Tyne, UK.)

(A)



(B)

Marker	Alleles	Allele Length	Peak Area
D13S628	2	454.4:461.8	7680:7262
D13S634	2	391.4:406.3	7003:7181
D18S1002	2	341.8:354	6083:5638
D18S386	2	352.2:381.7	7291:5847
D21S1411	2	316.9:333.4	6558:6037
D21S1446	1	212.4	14090
DXYS218	1	239.5	7669
HPRT	1	282.0	6169
SRY	0		
TAF_9	3(2:1)	3:X	40238:18508
X22	1	218	6680

Figure 11.3 Detecting sex chromosome aneuploidies using QF-PCR. (A) Marker sets. Primer pairs are designed to amplify X-specific markers *(HPRT),* Y-specific markers *(SRY),* markers in the pseudoautosomal regions PAR1 or PAR2 (shared by the X and Y), and highly homologous sequences on the X and Y chromosomes, such as the amelogenin genes *AMELX* and *AMELY* in which a single set of primers can amplify both sequences (which can be differentiated because of small length differences due to insertion or deletion). To gauge the ratio of X chromosomes to autosomes, primers are used to amplify equivalent segments of the *TAF9B* gene on Xq and a highly related

pseudogene *TAF9BP1* on 3p. CEN, centromere. Data from a practical example. The interpretation would be monosomy X (Turner syndrome) on the basis of the absence of the SRY marker and ratio of 2:1 for the length variants from *TAF9BP1* in chromosome 2 and *TAF9* on the X chromosome. (Data courtesy of Jerome Evans, NHS Northern Genetics Service, Newcastle upon Tyne, UK.)

Sex chromosome aneuploidies

The copy number of our sex chromosomes is more variable than that of auto-somes, ranging from monosomy (45,X) to different types of trisomy, tetrasomy, and occasionally pentasomy. Identifying a monosomy using PCR might seem particularly challenging—how can 45,X be distinguished from 46,XX? However, counting the sex chromosomes is possible by using primer sets specific for the X or Y chromosome plus primer sets that simultaneously amplify conserved sequences on both sex chromosomes or on both the X and an autosome (**Figure 11.3**).

Noninvasive fetal aneuploidy screening

Screening for fetal aneuploidies has been made possible by high-throughput sequencing of fetal DNA in maternal plasma, an advanced form of noninvasive prenatal screening. Because there is only a small number of viable human aneuploidies, it is relatively easy to design a series of QF-PCR assays for this purpose. We describe recent major advances in this area in Section 11.4.

Detecting large-scale copy number variants using chromosome SNP microarray analysis

Standard chromosome-banding karyotyping will not detect deletions and duplications less than 6–10 Mb of DNA. However, once the human genome had been sequenced techniques rapidly followed that allow us to detect

smaller copy number variants. Chromosome microarray analysis uses oligonucleotide sequences from well-studied polymorphic loci across the genome to monitor polymorphism at hundreds of loci on each chromosome. We previously explained the principle underlying microarray hybridization in Figure 3.9 on p. 71 and accompanying text in Section 3.3.

The most widely used type of chromosome microarray analysis uses single nucleotide polymorphism (SNP) microarrays. Test DNA samples from individuals are hybridized to microarrays containing oligonucleotides representing the different alleles at each of many thousands of SNP loci across the genome. SNP arrays do not directly compare a patient's test sample with a control sample. Instead, the assay compares the dosage of the individual being tested at any given locus with the equivalent values in a database of SNP array results from control individuals. Readouts of the SNP profiles across individual chromosomes allow us to detect gains and losses of sequences across the genome.

Deletions can be identified simply because of absence of heterozygosity across the deleted area: each SNP in the deleted area should show just a single allele. For duplications, the ratios of alleles will vary: if we imagine a SNP locus with two alleles, say A and B, a normal heterozygote would be scored as AB (with equal representations of alleles A and B); however, in regions of partial trisomy, loci in which both alleles are evident might show skewed allele ratios and might appear as AAB or ABB instead of AB (twice the signal for one allele compared to the other).

As well as identifying large deletions and duplications, SNP microarrays can identify regions of heterodisomy and isodisomy. That is useful for some conditions such as Prader-Willi syndrome (where uniparental disomy can be quite common, even if less frequent than paternal 15q11-q12 deletions), as illustrated in the case study profiled in **Clinical Box 15**.

CLINICAL BOX 15 CHROMOSOME SNP ARRAY ANALYSIS TO IDENTIFY MIXED HETERODISOMY/ISODISOMY IN A PRADER-WILLI SYNDROME CASE

David was the second child of unrelated parents Claire (aged 40) and Mike (aged 41). He was born by elective Caesarian section for breech presentation in good condition and weighed 2.8 kg. On Day 2 there was concern because he had never cried and was not interested in feeding, and so he was dependent on nasogastric tube feeds. He was examined by a pediatric neurologist who found him to have severe truncal hypotonia but normal tendon reflexes. She suspected Prader Willi syndrome and requested DNA methylation studies of the PWS critical region at 15q11-q13, which showed absence of the paternal allele, thus confirming the diagnosis. Follow-up chromosome SNP array analysis did not show any clinically significant copy number changes and excluded a 15q11-13 deletion, but revealed two large regions of homozygosity on chromosome 15 (see **Figure 1**).

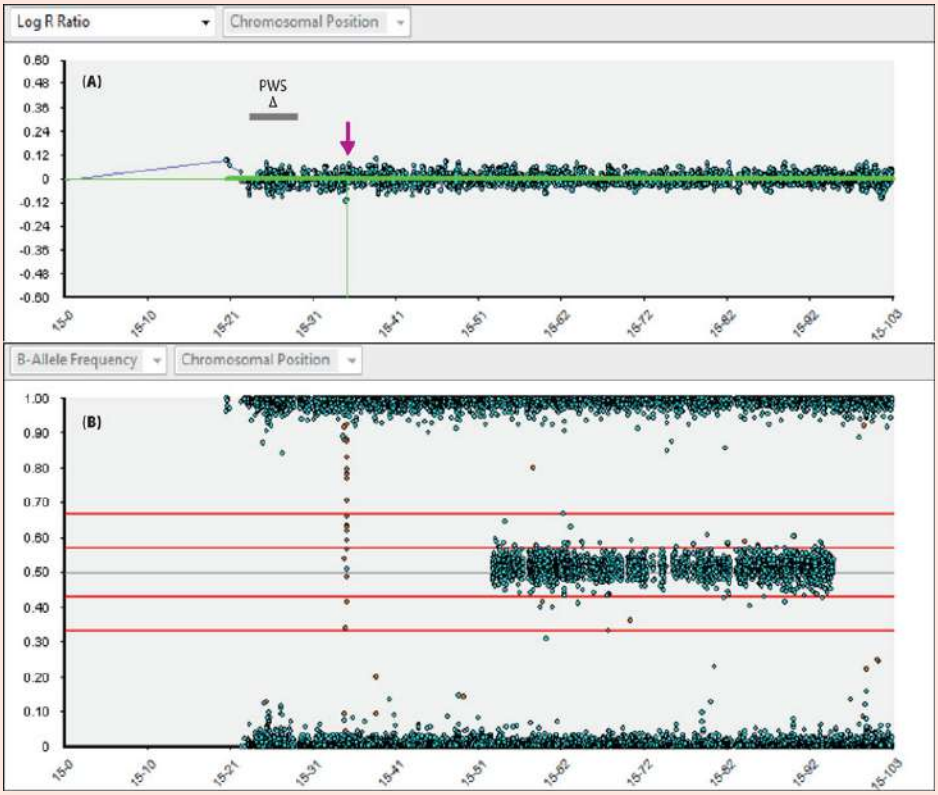


Figure 1 Chromosome SNP array analysis. The scale at the bottom of (A) and (B) represents the 103 Mb chromosome 15 from 15ptel (15-0) to 15qtel (15-103). (A) LogR ratio—the relative fluorescence of SNPs compared to SNPs with a diploid

complement. SNPs showing copy number gain would have a positive value; those showing deletions would have a negative score. Almost all SNPs, including over the PWS deletion area (PWS Δ) show a zero logR ratio, indicating two copies (the arrow marks a nullisomic loss where the logR ratio drops off the scale). (B) B-allele frequency—the ratio of alleles calculated by B/A+B. In a diploid situation the SNP data points should be arranged in three rows, with homozygous SNPs at top and bottom (B-allele frequency of 1 and 0, respectively) and a central row showing SNPs that are heterozygous. Here, the heterozygote row is confined to just a part of the chromosome (roughly 42 Mb of DNA from 15–52 to 15–94), with two flanking regions of segmental absence of heterozygosity.

The most likely explanation is that this case represents uniparental disomy with a mixture of heterodisomy for the central region of 15q (from the 15–52 to 15–95 positions) and isodisomy (for the flanking regions). The conceptus initially had trisomy 15 because a maternal egg with two different chromo-some 15s was fertilized by a normal sperm with a single chromosome 15. Non-disjunction of the two maternal chromosome 15s had occurred after a double crossover so that a central region of 15q was heterodisomic and flanked by isodisomic segments. Trisomy rescue ensued by loss of the paternal chromosome 15.

The family were seen by a clinical geneticist when David was three months old. He was by then able to be bottle fed and was gaining weight on high calorie feeds. He was still floppy but his head control was improving. David was referred to a pediatric endocrinologist and was treated with growth hormone from 10 months of age. At the age of 3 years his height and weight were on the 50th centile and by this stage he had not yet developed any food-seeking behavior. Maternal uniparental disomy for chromosome 15 accounts for 20–25 % of cases of PWS. It occurs due to rescue of a trisomy 15 conception and the recurrence risk in a future pregnancy is negligible.

Detecting and scanning for oncogenic fusion genes using, respectively, chromosome FISH and targeted RNA sequencing

Balanced translocations and inversions are not detectable by chromosome microarray analyses because there is no appreciable loss or gain of DNA. Both can, however, be detected by traditional chromosome-banding karyotyping, but their contribution to pathogenesis is quite different. Inversions and constitutional translocations are rare, and make a very small contribution to genetic disease. Somatic translocations, by contrast, are common in many kinds of cancer, creating oncogenic *fusion genes* thought to account for ~20 % of human cancer morbidity.

The high overall frequency of somatic translocation in cancer occurs because translocations provide an opportunity for inappropriate oncogene activation. If translocation breakpoints occur in the immediate neighborhood of a protooncogene on one chromosome and close to an actively transcribed gene on the other chromosome, the proto-oncogene can be brought into close proximity to active transcriptional control signals and be inappropriately expressed as part of a fusion gene—Figure 10.8 on page 379 shows the example of *ABL1* activation in chronic myelogenous leukemia after fusing with the *BCR* gene. Oncogenic fusion genes like this are thought to account for ~20 % of human cancer morbidity, but the prevalence of fusion genes shows significant differences in different cancers, and many fusion genes are specific to certain types of cancer.

Traditional chromosome-banding karyotyping in metaphase chromosome preparations takes a minimum of two weeks to deliver results: cells need to be grown in culture and then a spindle poison added during periods of active growth to arrest cells in metaphase. Modern alternative methods, however, can permit rapid detection of fusion genes in interphase cells, such as by using chromosome **fluorescence *in situ* hybridization** (FISH) and targeted RNA sequencing of fusion gene transcripts.

The nature of chromosome FISH and its applications

The essence of chromosome fluorescence *in situ* hybridization (FISH) is to fix either metaphase or interphase chromosome preparations on microscopic slides, treat the slides so as to denature the DNA, and hybridize fluorescently labeled probes of interest to the denatured DNA—see **Figure 11.4A**. Obtaining metaphase chromosome preparations from blood samples takes time (because of the need for cell culturing), and the locations of the fluorescent signals are typically recorded against a background stain that binds to all DNA sequences.

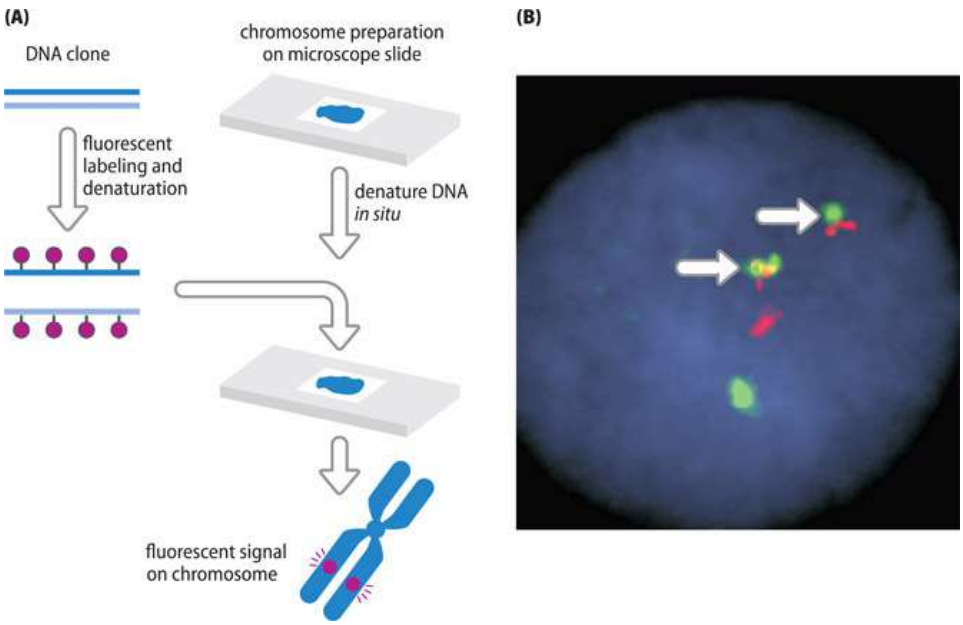


Figure 11.4 Principle of chromosome FISH and use in detecting recurrent translocations producing a fusion gene. (A) Principle of chromosome FISH. A labeled DNA clone of interest is hybridized to a denatured chromosome preparation on a microscope slide (either a metaphase chromosome preparation or an interphase chromosome preparation). When metaphase chromosome preparations are used, as shown here, a double fluorescent signal is often seen, representing hybridization to target sequences on the sister chromatids. (B) Interphase FISH to detect recurrent t(9;22) translocations in chronic myeloid leukemia (CML). Cases with CML often show translocations with breakpoints in the *ABL1* gene (on 9q) and the *BCR* gene (on 22q; see Figure 10.8A). Here, the *ABL1* and *BCR* probes give, respectively, red and green fluorescent signals. The white arrows indicate characteristic color signals for the fusion genes: very close positioning of red and green signals, with sometimes overlapping

orange-yellow signals. By contrast, the red and green signals at bottom are well separated and represent the normal chromosome 9 and normal chromosome 22, respectively. (Courtesy of Fiona Harding, Northern Genetics Service, Newcastle upon Tyne, UK.)

Chromosome FISH is often used to confirm regions of chromosome duplication or deletion that have been suggested by primary screening methods, notably chromosome microarray analysis. It can also be used to rapidly detect in interphase cells two types of oncogene activation in cancer:

- *gene amplification* in particular types of cancer, notably amplification of the *MYCN* gene in neuroblastoma—see Figure 10.7A on page 377)
- *fusion genes* generated by recurrent translocations. Chromosome FISH is notably used for detecting fusion genes in certain types of cancer that are strongly associated with recurrent translocations producing a specific fusion gene. For example, translocations associated with promyelocytic leukemia always produce *PML-RARA* fusion genes, and translocations producing *BCR-ABL1* fusion genes are common in chronic myeloid leukemia.

To detect specific fusion genes, interphase FISH can be carried out using probes representing the two genes that form the fusion gene (selected from regions retained on the translocation chromosomes). In the absence of visible chromosomes, probes from the two genes are labeled with different fluorophores, so that one produces a red fluorescent signal, for example, and the other produces a green fluorescent signal. The translocation chromosomes can readily be identified because here the green and red fluorescent signals are superimposed (Figure 11.4B).

Targeted RNA sequencing of fusion gene transcripts

The application of chromosome FISH to assaying oncogenic fusion genes arising from translocations is limited because the tests typically assay for the presence of a single fusion gene. They are therefore suited to detecting fusion genes frequently found in certain types of cancer, such as acute promyelocytic leukemia, chronic myeloid leukemia, and promyelocytic leukemia. Using chromosome FISH to carry out general screens for oncogenic fusion genes across a range of cancers would be very time-consuming and costly.

There is a considerable need for methods that can rapidly and simultaneously screen for a wide variety of fusion genes in cancers. And there can be considerable diagnostic value. For example, cells in different types of sarcoma appear very similar ("small blue round cells"), but different types of carcinoma—such as Ewing sarcoma, rhabdomyosarcoma and synovial sarcoma—are associated with particular types of recurrent translocation and specific oncogenic *fusion genes* that can therefore act as diagnostic aids.

In principle, RNA sequencing in cancer cells could allow genome-wide screens for fusion genes (in RNA sequencing RNA transcripts are first converted into double-stranded cDNA then sequenced). However, standard RNA sequencing does not have sufficient sensitivity to detect low expression signals from fusion genes (against the voluminous background expression signals from complex whole transcriptomes). To overcome the difficulty, *targeted RNA sequencing*has been developed using panels of specific biotinylated oligonucleotide probes to selectively enrich for sequences from RNA transcripts of interest. The same principle is used in the *targeted DNA sequencing* that selectively enriches for panels of gene sequences to screen for mutations; we will describe the method in detail in Section 11.3 (Box 11.2) when we consider methods of scanning for point mutations.

The current leader in commercial targeted RNA sequencing of fusion gene transcripts is the Illumina TruSight Pan-Cancer panel. It targets a total of 1385 cancer genes for gene expression, variant and fusion detection,

including detection of gene fusions with both known and novel gene fusion partners.

Detecting pathogenic moderate-to small-scale deletions and duplications at defined loci is often achieved using the MLPA or ddPCR methods

Structural variation includes moderate- to small-scale copy number variants (CNVs) where the sequences are >50 bp in length but below the limit of detection of chromosome banding-karyotyping, and often from hundreds of nucleotides to tens of kb in length. Two important classes of CNVs of this size range are listed below.

- *Inactivating intragenic deletions and duplications*. Large genes are often prone to intragenic deletions and duplications (often resulting from inappropriate pairing of repeat sequences. Pathogenesis may result from loss of important sequence (caused by deletions) or frameshifts (after deletion or duplication of one or more exons). Such copy number variation is typically assayed by multiplex ligation-dependent probe amplification (MLPA) as described below.
- *Unstable oligonucleotide repeat expansion*. PCR assays are often used to amplify the region containing the expanded repeat. The primers are labeled with a 5' fluorophore that gets incorporated into the product that can be detected after capillary electrophoresis. For very large expansions, however, more specialized triplet repeat-primed PCR assays are used— see below.

Various laboratory methods permit analysis of large CNVs. The versatile MLPA method is used widely to scan for exon deletions and duplications, but is limited by availability of commercial kits, as explained below. An alternative method that is rising in importance is a type of quantitative PCR known as droplet digital PCR (ddPCR). We give detailed descriptions of MLPA and ddPCR later in this section. Two additional methods are used infrequently, and we describe them briefly immediately below.

Southern blot-hyridization is very rarely used now, although it continues to be used for identifying pathogenic large deletions in facioscapulohumeral dystrophy, as shown in Clinical Box 3 on page 170, and sometimes for very large expansions of unstable short tandem repeats. It involves digesting genomic DNA with suitable restriction nucleases, size-fractionating the resulting fragments by agarose gel electrophoresis, denaturing the DNA *in situ* in the gel, then transferring the DNA to a plastic sheet laid over the gel so that the fragments are located on the plastic in a mirror image representation of how they appeared on the gel. Thereafter, the denatured DNA on the filter is hybridized with a suitable labeled gene probe. Readers interested in the details can find an adequate explanation of the technique in the online encyclopedia Wikipedia.

Triplet repeat-primed PCR is a very specialized technique used to detect particularly large-scale expansion of unstable short tandem repeats, as may be seen in disorders such as myotonic dystrophy and Fragile X syndrome. The output of an application in a myotonic dystrophy case study is shown in Clinical Box 5 on page 196. Readers interested in the details of the technique can find them at PMID 9004136.

Multiplex ligation-dependent probe amplification (MLPA)

MLPA is a quick and versatile method that can detect copy number changes over a broad range of DNA lengths. It uses pairs of short single-stranded sequences (called *probes*, but quite distinct from hybridization probes). The probes are designed to bind to specific exons (or to other target sequences whose relative copy number we wish to determine). Each pair of probes is designed to hybridize collectively to a *continuous* target DNA sequence; that is, when they bind to the target DNA, the pair of probes align immediately next to each other. The gap between them can then be sealed using DNA ligase to give a single probe that is complementary in sequence to the target sequence of interest (**Figure 11.5**).

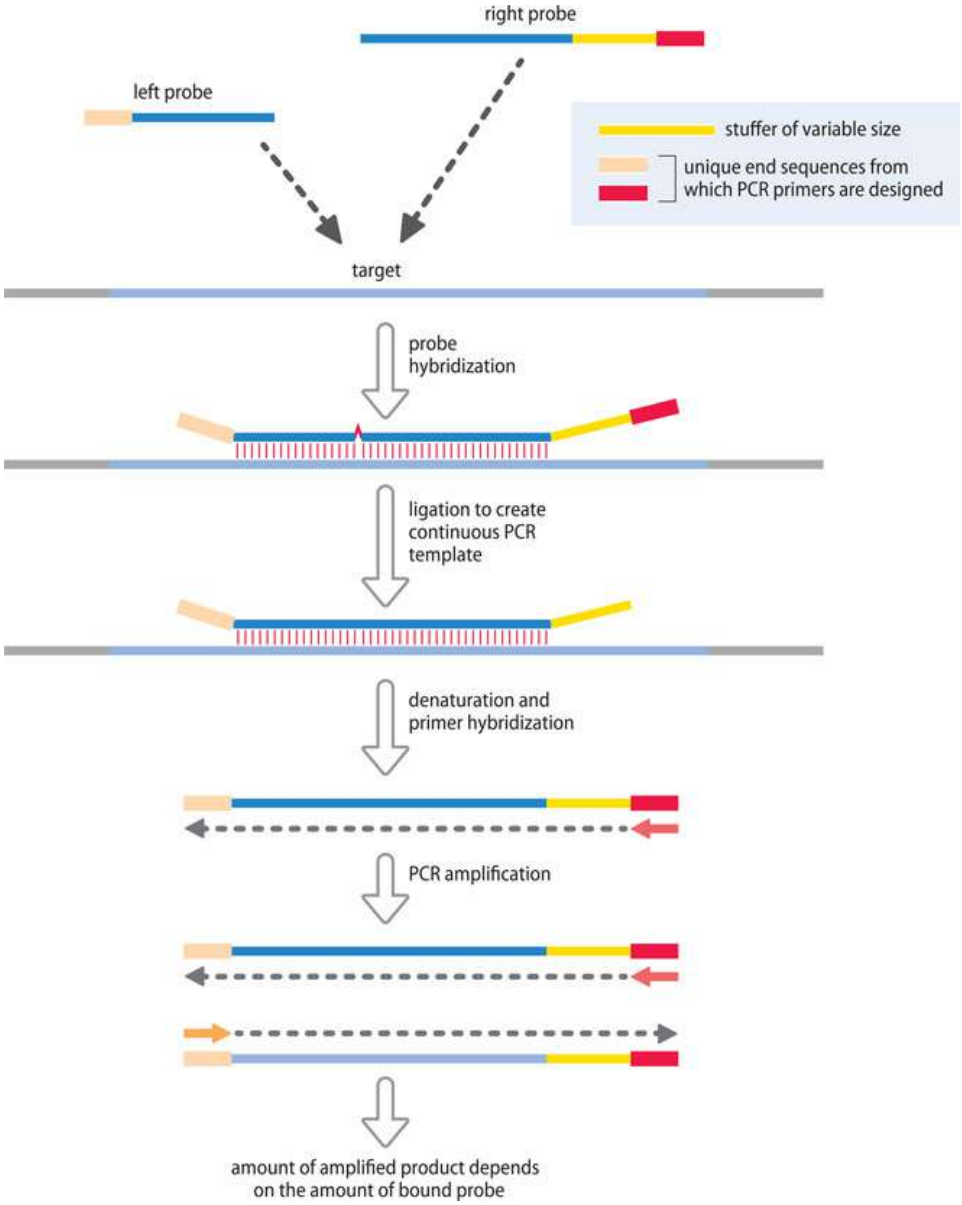


Figure 11.5 The principle of multiplex ligation probe amplification (MLPA). For each target sequence (such as an individual exon), a pair of probes is designed that will hybridize to adjacent sequences within the target and will carry unique end sequences not present in the genome. The aim is to use DNA ligase to seal the left and right probes to give a continuous sequence flanked by the unique end sequences and then to amplify the continuous sequence by using primers complementary in sequence to the unique end sequences. Probe pairs for multiple different target sequences (such as multiple exons within a gene) are simultaneously hybridized to their target sequences and ligated to form continuous sequences that are then simultaneously amplified in a multiplex reaction. The point of the stuffer fragment is simply to provide a spacer sequence whose

length can be varied. This can ensure different sizes of the PCR products from a multiplex reaction (in which multiple probe sets are used simultaneously) so that the products can be readily separated by capillary gel electrophoresis.

The 5¢ end of one of the probe pair and the 3¢ end of the other are designed to contain unique sequences not present in our genome. By designing oligonucleotide primers that will bind to regions in the unique end sequences only, the probe sequences are selectively amplified in a PCR reaction.

A key feature of MLPA is that the amount of amplified probe product is proportional to the number of bound copies of the probe, which in turn depends on the number of target sequences the probe has bound to. With a heterozygous deletion, for example, there is one copy of the target sequence instead of two; the amount of bound (and therefore ligated) probe is one-half of the normal amount, and the amount of amplified product is proportionally reduced.

Often, MLPA is designed to be a *multiplex* reaction: up to 55 pairs of probes can be used to bind simultaneously to different target sequences. The left and right probes for each target sequence all have the same set of left and right unique end sequences, and so all ligated probes can be amplified by a common set of primers (specific for the unique end sequences). But the stuffer sequences (described in Figure 11.5) are designed to be of different lengths for different probes, enabling the amplified probes to be physically separated by capillary gel electrophoresis and quantified independently. **Figure 11.6** gives a practical example of how MLPA can be used to screen simultaneously for large numbers of different exons. Useful YouTube video and text summaries of MLPA technology can be accessed from the MRC-Holland website at https://www.mrcholland.com/technology/mlpa/technique

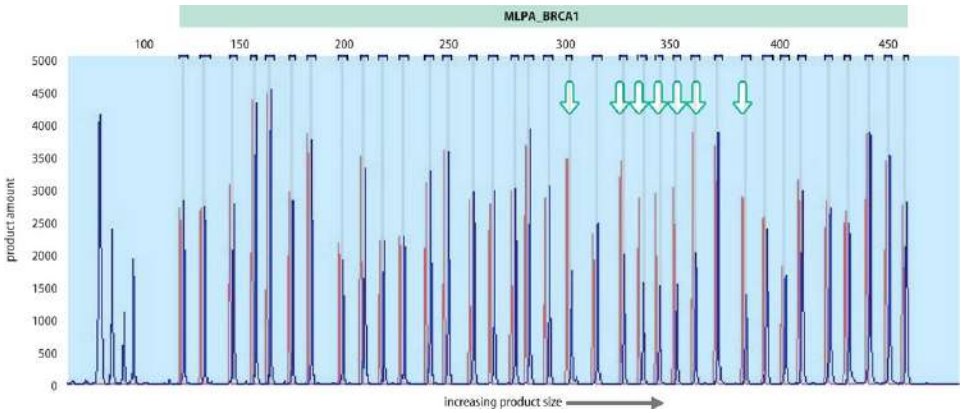


Figure 11.6 Using MLPA to scan for constitutional copy number changes in the exons of the *BRCA1* gene. MLPA scan. The blue peaks at the left from 0 to 110 bp on the horizontal size scale are internal controls. The paired blue and red peaks in the size range 125–475 bp represent comparative MLPA results in a normal control sample (blue) and a test sample (red) for individual exons of the *BRCA1* gene in most cases (however, for some large exons two partly overlapping probes were used). The test sample came from an individual with breast cancer in whom previous DNA sequencing investigations were unable to identify changes in the exons of the *BRCA1* gene. The MLPA analysis shown here identified a deletion that encompassed seven *consecutive* exons (marked by vertical green-outlined arrows)—in each case the blue peak is reduced by roughly one-half, as expected for a heterozygous deletion. Note: the order of the peaks is not the same as the order of the exons within the gene. (Data courtesy of Louise Stanley, NHS Northern Genetics Service, Newcastle upon Tyne, UK).

Droplet digital PCR (ddPCR)

Digital PCR (dPCR), like real-time PCR is a type of quantitative PCR requiring a specialist PCR machine. It offers precise absolute quantification (unlike real-time PCR, which requires reference standards to permit quantitation). A key feature is that individual DNA samples are extensively divided into very many smaller samples, each with limiting amounts of template DNA but with all the components required for PCR amplification of any template DNA. Like real-time PCR, it uses fluorescent probes.

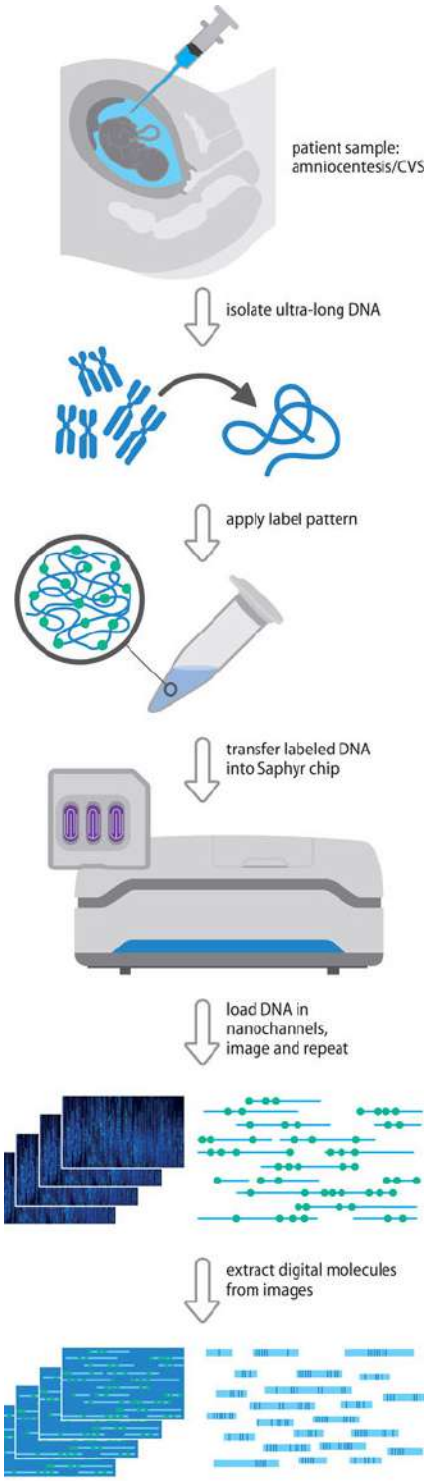
In the popular droplet digital PCR (ddPCR) method, individual DNA in a standard PCR reaction mix are mixed with oil to create an emulsion and microfluidic devices are used to divide a single PCR reaction sample into 20 000 droplets that are individually distributed into microwells containing either zero or one (or at most a very few) template DNA molecules. Each droplet is effectively a PCR mini-reactor where PCR amplification and analysis occur separately from all the other droplets. After the reaction is over, the droplets are individually counted and scored as positive or negative for fluorescence, with application of Poisson statistical data analysis to enable highly accurate DNA quantification. A YouTube video explaining the method is available at https://www.youtube.com/watch?v=Qwma-1Ek-Y4&ab_channel=Bio-RadLaboratories

Two very different routes towards universal genome-wide screens for structural variation: genome-wide sequencing and optical genome mapping

As technology develops rapidly, so too does change. The current profusion of techniques for detecting structural variation may soon be replaced by universal screening systems that detect all types of structural variation. Currently, there are two contenders.

- *Whole genome sequencing* (*WGS*). Because WGS is also used to identify point mutations we defer considering this until we cover genetic testing of point mutations in Section 11.3. Suffice it to say that it can also detect structural abnormalities across the genome.
- *Optical genome mapping.* This is a very new and radically different way of scanning structural variation across the genome. Extremely long DNA fragments are first isolated from sample cells, labeled and imaged. Individual color images are then aligned with equivalent patterns from reference genomes to identify structural variation.

A pioneer in optical genome mapping is the Bionano company, whose Saphyr system has recently been released and is being adopted by many health service laboratories. See **Figure 11.7** for an overview of the Saphyr system.



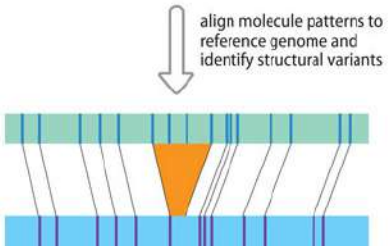


Figure 11.7 Prenatal workflow for optical genome mapping using the Bionano Saphyr system. DNA obtained from cells contained in amniotic fluid or a chorionic villus sample is labeled across the genome at a 6 bp motif by a labeling technology; the resulting label pattern is unique to each individual sample. Labeled ultra-long DNA molecules are loaded on a Saphyr chip, then electrophoresed into nanochannels where they are uniformly linearized for imaging in repeated cycles. The resulting images are processed to extract molecules containing the linear positions of sequence motif labels. Multiple molecules are used to create consensus genome maps representing different alleles from the sample. The sample's unique optical genome map is aligned to the reference genome and differences are automatically called, allowing for detection of structural variations in a genome-wide fashion. (Image modified from: https://bionanogenomics.com)

11.3 GENETIC AND GENOMIC TESTING FOR PATHOGENIC POINT MUTATIONS AND DNA METHYLATION TESTING

The most common pathogenic DNA changes are point mutations, mostly single nucleotide changes, or changes to a very small number of nucleotides. These changes can be detected by a range of different methods. Some long-established methods are designed to detect a specific mutation in a defined gene or to scan for any point mutation in a gene or panel of genes associated with a specific disease.

More recently, improvements in the speed and cost of DNA sequencing have brought radical changes to clinical settings. Instead of starting with a pheno-type, or family history of a phenotype, and considering which gene(s) should be analyzed, it is now often easier to perform genome-wide

mutation scanning by whole exome sequencing or whole genome sequencing, and then apply the filtering at a later stage. **Next-generation sequencing (NGS)**, which encompasses a variety of sequencing technologies that use massively parallel DNA sequencing has driven that change (see **Box 11.1** for an overview). Sanger (dideoxy) sequencing, previously described in Figure 3.10 on page 72, remains in use, but typically as a confirmatory technique.

BOX 11.1 MASSIVELY PARALLEL ("NEXT-GENERATION") WHOLE-GENOME SEQUENCING

In standard dideoxy sequencing, individual DNA sequences of interest must first be purified; they are then sequenced, one after another. The sequencing involves DNA synthesis reactions, producing a series of reaction products of different lengths that are then separated by capillary gel electrophoresis (Figure 3.10, page 72). By contrast, massively parallel DNA sequencing (often called next-generation sequencing) is indiscriminate: all of the different DNA fragments in a complex starting DNA sample can be simultaneously sequenced without any need for gel electrophoresis. That allows a vastly greater sequencing output

There are many different types of massively parallel DNA sequencing, but they can be separated into two broad categories: those in which the starting DNA sequences are first amplified by PCR, and those that involve single molecule sequencing (that is, sequencing of unamplified DNA molecules). We give details of run parameters for major commercially available technologies in Table 3.3 on page 75.

Massively parallel DNA sequencing often involves *sequencing-by-synthesis*. That is, the sequencing reaction is monitored as each consecutive nucleotide is inserted during DNA synthesis. **Figure 1** shows the workflow that is involved in massively parallel sequencing and gives a simplified illustration of a popular form of sequencing-by-synthesis used by the Illumina company. Alternative methods are used by some other companies.

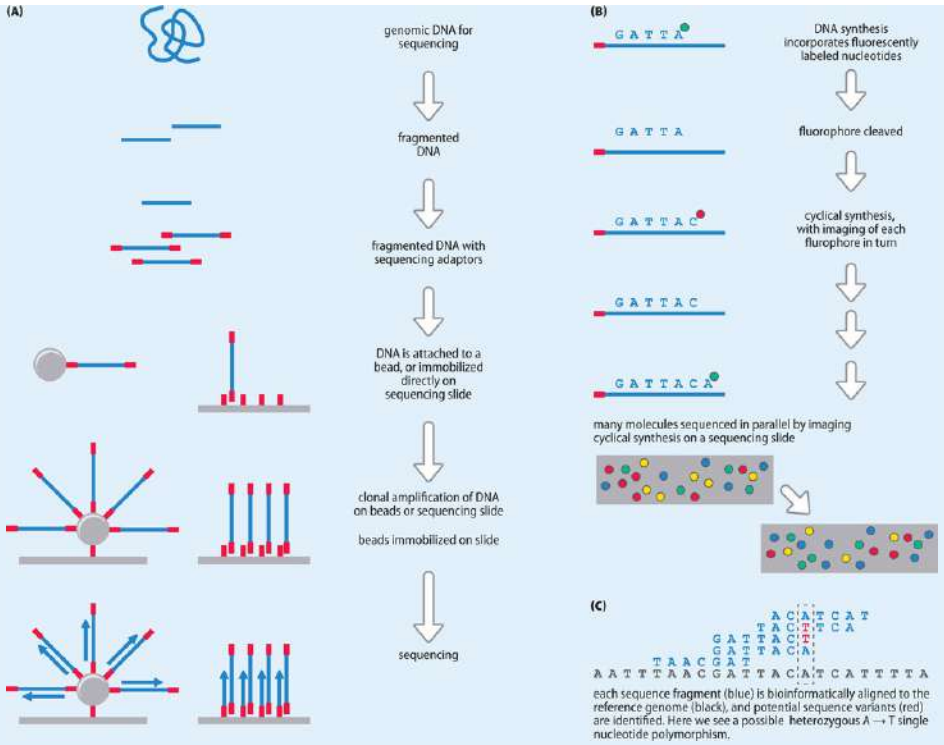


Figure 1 Next-generation sequencing workflow. (A) Genomic DNA is fragmented, and adaptor oligonucleotides are attached. The DNA is then attached either to a bead or directly to the sequencing slide. In either case, the DNA is clonally amplified in this location to provide a cluster of molecules with identical sequences. If beads are used, they are then immobilized on a sequencing slide. (B) The Illumina Genome Analyzer system of sequencing-by-synthesis. The sequence of each fragment is read by decoding the sequence of fluorophores imaged at each physical position on a sequencing slide. Advanced optics permit massively parallel sequencing. (C) Each DNA fragment yields one or two end sequences, depending on whether it is sequenced from one or both ends. These end sequences are computationally aligned with a reference sequence and mismatches are identified. (From Ware JS et al. [2012] *Heart* 98:276–281; PMID 22128206. With permission from the BMJ Publishing Group Ltd.)

We begin with a description of the range of techniques that can be used to identify specific *defined* pathogenic single nucleotide variants (SNVs). Then we consider how one can *scan* for undetermined pathogenic point

mutations, generally by using DNA sequencing at different levels. That started with standard Sanger dideoxy sequencing across exons and exon-intron boundaries plus promoter regions of a specific gene, but often now includes "gene panels", amplified sequences from multiple genes that are obtained by *targeted DNA sequencing*. At a higher level, genome-wide scanning has begun to be used quite widely, beginning with whole exome sequencing, and more recently whole genome sequencing.

Thereafter, we move on to consider the problem of sequence *interpretation.* Sometimes trying to identify a pathogenic SNV in a single gene can be difficult, but the problem scales up massively when genome-wide sequencing is undertaken: there are huge numbers of candidate pathogenic variants to sift through, as described below. Although a battery of online resources can help us weed out less compelling variants, the task is nevertheless complex, especially for whole genome sequencing. Not only will many variants of uncertain significance arise in the sequence, but there are also ethical issues relating to the discovery of incidental findings that may raise information about additional health issues beyond the health concern for which the test was ordered.

We end this section with testing for DNA methylation changes. Such changes are especially important in cancers, but are also relevant to certain single gene disorders, notably those associated with imprinting defects.

Diverse methods permit rapid genotyping of specific point mutations

Instead of simply scanning for potentially pathogenic mutations, genetic tests may seek to identify a *specific* point mutation at a defined locus. That may be required for different reasons. A member of a family may be found to have a specific pathogenic mutation and there is interest in knowing if other members of the family have the same mutation, or if the mutation has been inherited in early pregnancy. At the population level there may be interest in screening carriers of common mutations such as the sickle cell mutation. And in cancers, tumor biopsies can be tested for the presence of

specific causative mutations to monitor minimal residual disease post-treatment, and to check for mutations that would govern the tumor's response to a targeted drug. DNA sequencing can be used to identify such variants, but it is often much more convenient to use simpler alternative detection methods.

The vast majority of pathogenic point mutations are SNVs, often occurring by substitution, but sometimes by insertion or deletion. Different methods can discriminate between the mutant and normal alleles. For mutations like these, a pair of allele-specific oligonucleotides (ASO) can be designed that represent mutant or normal allele sequences encompassing the sequence containing the point mutation, so that a mutant-specific ASO base pairs perfectly with the mutant sequence and a normal sequence-specific ASO base pairs with the corresponding normal allele. Normal and mutant ASOs that differ at a single base at a central position in the oligonucleotide sequence may allow allele-specific hybridization and that is the basis of SNP-chip hybridization that we described in Section 11.2 and in GWA studies in Chapter 8.

Microarray-based hybridization is very rarely used now for genotyping pathogenic point mutations. Instead, the genotyping quite often depends on designing the mutant ASO and normal ASO to have a single base difference at the 3¢-terminal nucleotide, corresponding to the position and identity of the single-nucleotide change to be tested. The mutant ASO base pairs perfectly with the mutant allele; the normal ASO is specific for the normal allele sequence. When base pairing occurs between the normal ASO and mutant DNA, however, there is base mismatch at the 3¢ end nucleotide of the ASO (see **Figure 11.8A**); the same applies in the case of the mutant ASO binding to normal DNA.

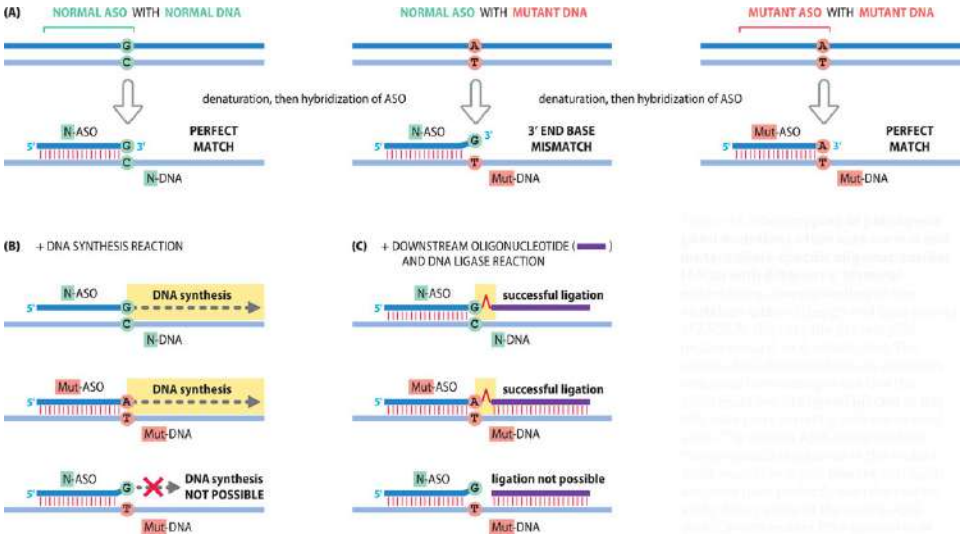


Figure 11.8 Genotyping of pathogenic point mutations often uses normal and mutant allele-specific oligonucleotides (ASOs) with different 3′ terminal nucleotides, corresponding to the mutation site. (A) Design and base pairing of ASOs. In this case the pathological mutation is a G ® A substitution. The normal ASO, designed from an upstream sequence terminating at the G of the point mutation site (green bracket at top left) base pairs perfectly with the normal allele. The mutant ASO, designed from the comparable sequence in the mutant allele ends in an A (red bracket, top right), and base pairs perfectly with the mutant allele. Base pairing of the normal ASO (N-ASO) with mutant DNA (center) or of mutant ASO with the normal allele (not shown) results in a 3¢ end base mismatch. (B,C) Exploiting the base mismatch using follow-up DNA synthesis (B) or DNA ligation (C). A normal ASO perfectly base-paired to normal DNA or a mutant ASO perfectly base paired to mutant DNA permits DNA synthesis using the ASO as a primer; but if there a 3¢ end base mismatch, DNA synthesis cannot be primed by the ASO. Similarly, an ASO perfectly base paired to the DNA can be ligated (indicated by the highlighted red chevron) to a downstream oligonucleotide that is base paired to an immediately adjacent downstream sequence, but an ASO with a 3¢ end base mismatch cannot be so ligated.

The difference in base pairing—whether an ASO is perfectly base paired or has a base mispaired—can be exploited in different ways. The *oligonucleotide ligation assay* exploits the base mismatch by following up with a DNA ligation reaction, one that assays the ability to ligate the bound ASO to another oligo-nucleotide designed to base pair to an immediately

adjacent sequence on the template DNA (Figure 11.8C shows the principle). This type of assay is used less frequently now. Instead, some of the more frequently used methods exploit base mismatching (sometimes at the 3¢ end and sometimes a central base mismatch) with a subsequent DNA synthesis reaction (Figure 11.8B). The methods listed below are three of the more commonly used for genotyping SNVs. In addition, mass spectrometry is used for genotyping and we describe that in the section on multiplex genotyping that follows this one.

- *Amplification refractory mutation system* (*ARMS*). The ARMS method exploits the inability of an ASO with a 3¢ end-base mismatch to prime DNA synthesis in a PCR reaction (see **Figure 11.9**).

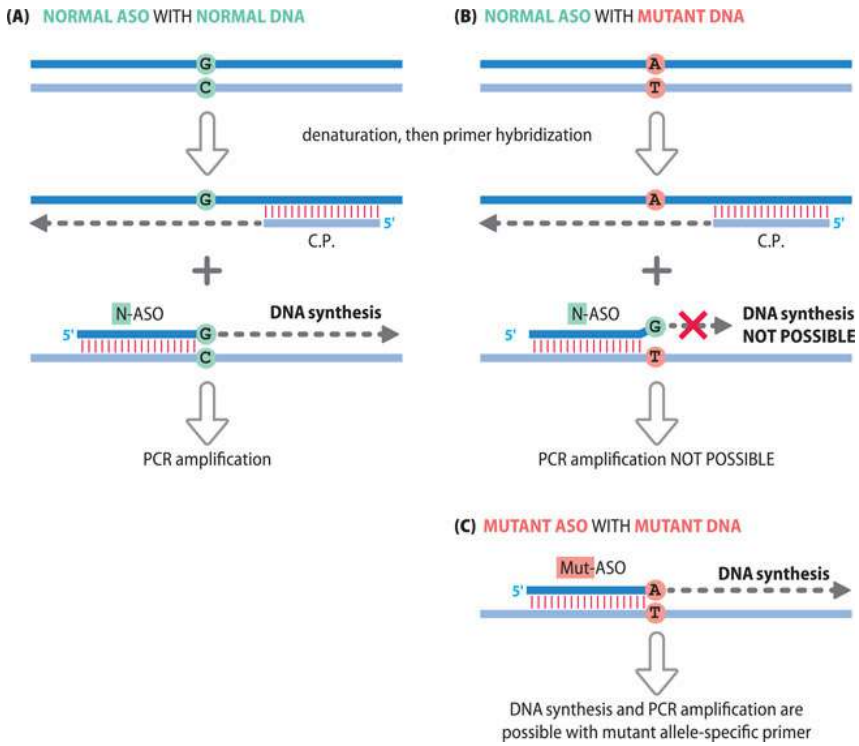


Figure 11.9 The principle underlying the amplification refractory mutation system (ARMS). In a PCR assay, when using a DNA polymerase lacking a proof-reading activity, the nucleotide at the 3¢ end of each oligonucleotide primer needs to be correctly base paired to the template DNA to allow DNA synthesis. (A) The normal allele-specific oligonucleotide

primer (N-ASO) terminating in a G at its 3¢ end will allow DNA synthesis (and consequent PCR amplification) because it is correctly base paired to the normal template DNA. (B) The mutant template DNA has a mutation with a single G to A difference. The N-ASO has its 3¢ terminal nucleotide at the nucleotide position corresponding to the mutation site. The lack of base pairing at the 3¢ end means that DNA synthesis cannot be primed. The same is true for mutant ASO with normal DNA (not shown). (C) However, if a mutant allele-specific oligonucleotide primer (Mut-ASO), terminating in an A at its 3¢ end, is used with mutant DNA, DNA synthesis can occur and PCR amplification is possible. Thus, the primer terminating in a G would be specific for normal alleles, and the primer terminating in an A would be specific for the mutant allele. C.P., common primer.

- *Real-time PCR using TaqMan genotyping.* An ASO with a central base mismatch is inadequate in priming DNA synthesis in a quantitative PCR reaction where amplification is tracked using a quantitative fluorescent signal (see **Figure 11.10**).

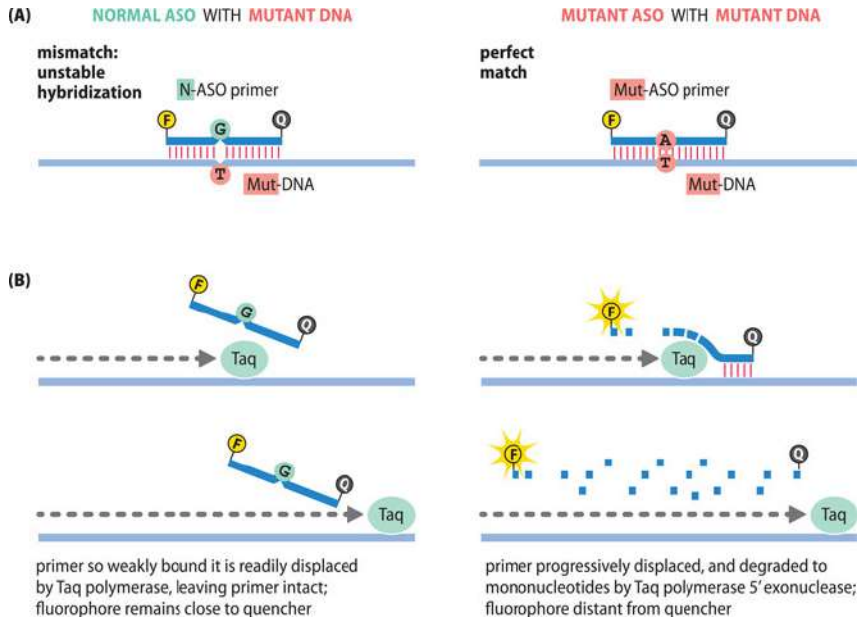


Figure 11.10 Genotyping point mutations using real-time PCR with the TaqMan™ double-dye system. (A) Primer hybridization. A central base mismatch between, for example, a normal allele-specific oligonucleotide (N-

ASO) primer and the complementary mutant DNA strand, results in unstable hybridization, unlike a perfectly base paired ASO. (B) DNA synthesis. In real-time PCR with TaqMan^{TM} genotyping, DNA is synthesized with *Taq* polymerase and monitored continuously by fluorescence. That is possible because the ASO primers carry two different dyes: a 5¢ fluorophore (F) and a 3¢ quencher (Q). The fluorophore can emit strong fluorescence but when in close range to the quencher its fluorescence signal is suppressed. When a DNA synthesis step occurs during PCR, the Taq polymerase effortless displaces a weakly bound ASO primer with a central base mismatch, which is simply brushed aside. But a perfectly base paired primer is displaced progressively by Taq polymerase whose associated 5¢®3¢ exonuclease activity sequentially degrades the ASO into mononucleotides that are dispersed in solution, thereby liberating the fluorophore from the inhibitory quencher to cause fluorescence.

- *Pyrosequencing*. Exploits the inability of an ASO with a 3¢ end-base mismatch to prime DNA synthesis in a single DNA synthesis step. The method is so-called because, just as in a sequencing reaction, it follows incorporation of an individual nucleotide in a growing DNA chain through the reaction: dNTP®dNMP + PPi where PPi represents the pyrophosphate released from the dNTP to enable incorporation of a dNMP. The released pyrophosphate is used to drive a color reaction (see **Figure 11.11**).

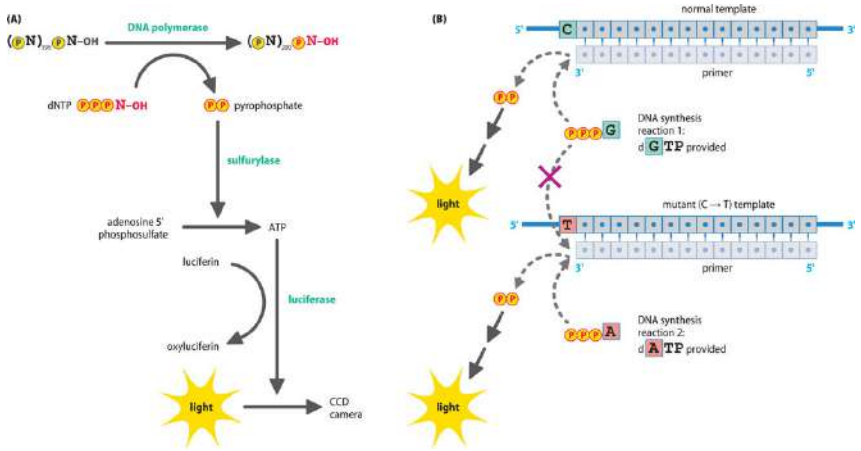


Figure 11.11 The principle underlying pyrosequencing. (A) Incorporating a nucleotide into a growing DNA chain requires cleavage of the dNTP precursor and insertion of a dNMP residue. The remaining pyrophosphate is detected in pyrosequencing by a two-step reaction. First, ATP sulfurylase quantitatively converts pyrophosphate (PPi) to ATP in the presence of adenosine 5¢ phosphosulfate. Then the released ATP drives a reaction where luciferase converts luciferin to oxyluciferin, thereby generating visible light in amounts proportional to the amount of ATP. Each time a nucleotide is incorporated, a light signal is produced and recorded by a charge-coupled device (CCD) camera. (B) The dNTP precursors for DNA synthesis are provided *individually* and in a set order. If in the first reaction dGTP is the only nucleotide precursor, a G would be incorporated opposite the highlighted C in the normal template at the top, and a PPi residue would be released and trigger light production as shown in (A). But if a mutant DNA template were used (with T replacing C), no light would be produced in the first reaction (a G would not be inserted opposite the T residue). If, however, dATP is provided in a second reaction, an A would be inserted opposite the T in the mutant template, producing PPi and light, but no base incorporation would occur opposite the C in the normal template.

The advantages of multiplex genotyping

Some types of genetic disorder show very limited mutational heterogeneity. Sickle-cell anemia is the outstanding example: its very specific phenotype is always due to a nucleotide substitution that replaces a valine residue at position 6 in the b-globin chain by glutamate. Unstable oligonucleotide repeat disorders such as Huntington disease typically show a very limited range of mutations.

For most single-gene disorders, disease can be caused by any number of different causative point mutations. Nevertheless, certain pathogenic point mutations may be frequent in certain populations and contribute very significantly to disease. The *CFTR* mutation causing the p.Phe508 del cystic fibrosis variant is very common in populations of European origin,

for example, as are two hemochromatosis-causing variants in the *HFE* gene that result in the C282Y and H63D amino acid substitutions.

Multiplex genotyping can be performed with the genotyping methods described in the previous section. Generally, that means genotyping dozens of point mutations at a time and so that can be used as a type of *mutation scanning* to see if the pathogenic mutation under investigation is one among a set of *known* pathogenic variants. Commercial kits based on the ARMS method have permitted testing for 50 common cystic fibrosis-associated mutations, covering ~90 % of the pathogenic mutations found in some populations of north European ancestry that have a high frequency of cystic fibrosis. Pyrosequencing and real-time PCR with TaqMan genotyping can also be used in multiplex genotyping.

Multiplex genotyping using mass spectrometry

In mass spectrometry, samples are ionized into charged molecules and the ratio of their mass to charge is measured. In MALDI-TOF mass spectrometry, the ionization occurs by **m**atrix-**a**ssisted **l**aser **d**esorption/**i**onization and the mass analyzer is a **t**ime-**o**f-**f**light analyzer. Genotyping by MALDI-TOF mass spectrometry typically uses the single-primer extension method to add a single chain-terminating nucleotide that discriminates between mutant and normal alleles. To do this, a primer is annealed to denatured template DNA so that it binds to a sequence terminating one nucleotide upstream of the SNV (single nucleotide variant) site. In the presence of DNA polymerase and chain-terminating dideoxy(dd)NTPs the primer is extended by one nucleotide. Mass analysis then permits genotyping by discriminating between the different reaction products according to their mass (see **Figure 11.12**).

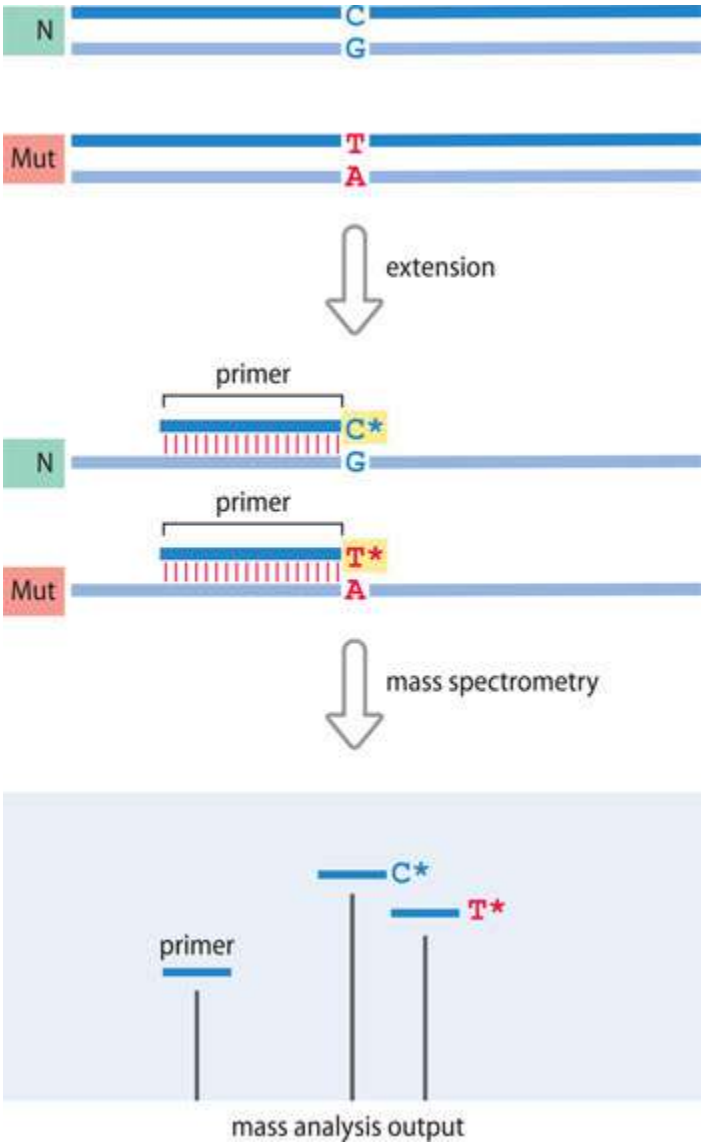


Figure 11.12 Genotyping single nucleotide variants (SNV) using mass spectrometry (MS). Imagine genotyping a heterozygote for a C>T substitution, using a primer designed to bind to the region immediately upstream of the SNV site. For the extension reaction, the DNA is denatured, and the primer, DNA polymerase and chain-terminating dideoxy(dd)NTPs are added, allowing incorporation of a *single* chain-terminating dideoxynucleotide, either ddC for the normal allele (shown as C*), or ddT for the mutant (shown as T*). On the basis of mass alone, the two reaction products, primer extended by C* and primer extended by T*, can be distinguished from each other and from the unextended primer.

Other genotyping methods can determine dozens of genotypes at a time, but MALDI-TOF mass spectrometry using a machine such as the Agena MassArray™ can genotype large numbers of variants. As a result, it has increasingly been used in diagnostic DNA services as an inexpensive way of carrying out genome-wide SNP analyses in *trio testing*, seeking to confirm biological relationships in two parents and affected child prior to carrying out more expensive whole exome or whole genome sequencing.

Mutation scanning: from genes and gene panels to whole exome and whole genome sequencing

As described in the previous section, mutation scanning may be carried out on a limited scale using multiplex genotyping of known pathogenic variants. But for the great majority of mutation scanning the object is to *define* a pathogenic variant whose identity may be difficult to suspect, and may never have been recorded previously. In some cases we may not even know the disease gene locus, and a genome-wide mutation scan may be needed, as described below. Quite often, however, we might wish to scan for mutations in a known disease gene, or in a *gene panel*, a group of genes associated with the same type of disorder. In that case, **targeted DNA sequencing** can be carried out: desired DNA sequences are captured from a genomic DNA sample by a DNA hybridization method and submitted for DNA sequencing. **Box 11.2** gives an overview.

BOX 11.2 TARGETED DNA SEQUENCING FOR MUTATION SCANNING

Targeted DNA sequencing means using a DNA hybridization method to capture desired DNA sequences from a genomic DNA sample so that they can be selectively sequenced, normally by massively parallel ("next-generation") sequencing. The capture method relies on the extraordinarily high affinity of streptavidin, a bacterial protein, for the vitamin biotin (see Figure 1).

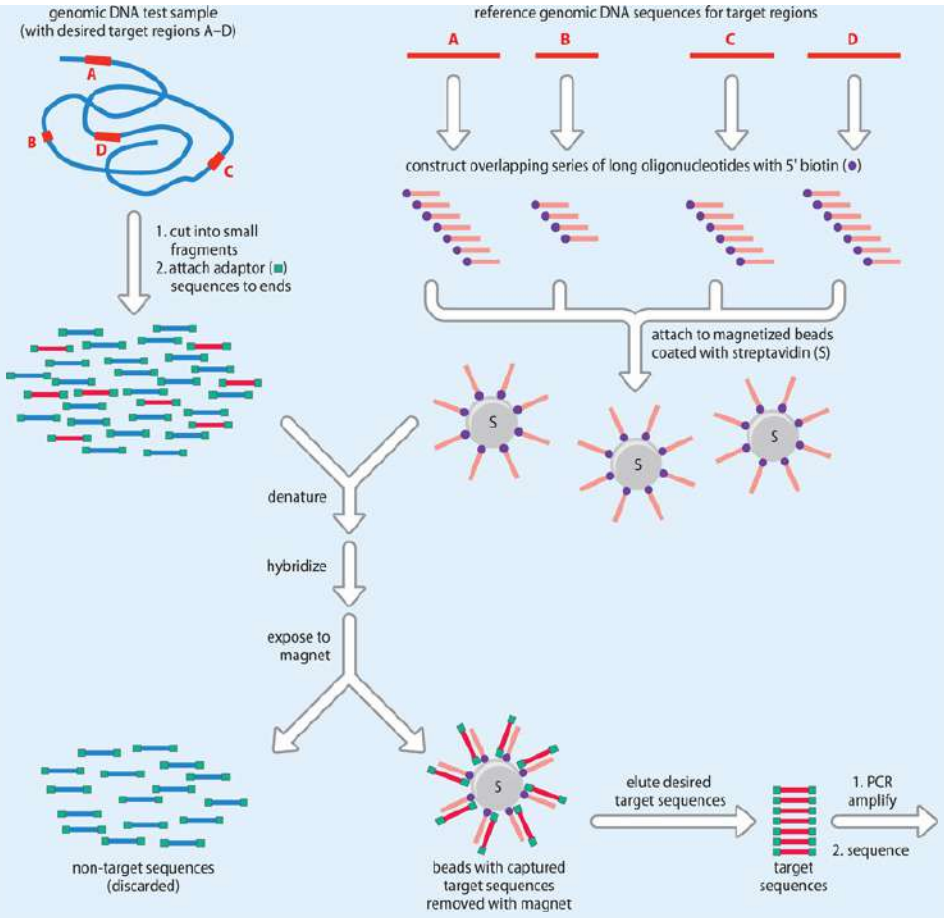


Figure 1 The principle of targeted DNA sequencing. In this example, four DNA regions, A–D, are targeted for sequencing. A series of partially overlapping oligonucleotides are synthesized to represent each of the desired target sequences (top right); each oligonucleotide has a biotin group covalently attached to its 5¢ end. The biotinylated oligonucleotides are then mixed with magnetized beads coated with streptavidin; the strong biotin-streptavidin affinity means that the oligonucleotides bind strongly to the beads. The genomic DNA sample is fragmented, denatured, and mixed with bead-oligonucleotide complexes. Sequences in the target regions of the genomic DNA sample (shown as red bars) will hybridize to complementary oligonucleotide sequences on the beads. Having fished out the target sequences the beads can be removed with a magnet, and the target sequences can be eluted, amplified, and sequenced.

Note: the target DNA sequencing method shown in Figure 1 can be adapted for **targeted RNA sequencing**. In RNA sequencing, RNA

transcripts isolated from cells are first converted to cDNA then sequenced. Targeted RNA sequencing involves converting total RNA to cDNA then capturing regions of interest using the method described in Figure 1. Targeted RNA sequencing is important for sequencing oncogenic fusion genes, as explained in Section 11.2.

TARGETED DNA SEQUENCING TO SCAN FOR MUTATIONS IN GENES AND GENE PANELS

Targeted DNA sequencing focuses on genes of interest. Much of the sequence of small genes and genes with small introns may be captured as overlapping sequences, but for a gene with many large introns, the focus can be on capturing exons and the immediately flanking intron sequences (to maximize detection of splice-site mutations) plus known major regulatory sequences.

Rather than scan individual genes, *gene panels* are commonly used now, in which sequences are captured from multiple genes that are often implicated in the same disorder, or a collection of similar disorders, and then submitted to massively parallel DNA sequencing. They often represent monogenic disorders and common cancers, and can be narrowly or broadly focused. **Table 1** provides some examples; comprehensive lists of curated gene panels are publicly available at sites such as the PanelApp database at https://panelapp.genomicsengland.co.uk and the GenCC database at https://thegencc.org/

TABLE 1 EXAMPLES OF HUMAN GENE PANELS USED IN TARGETED DNA SEQUENCING

Gene panel	Gene sequences
Breast cancer (common)	*BRCA1,BRCA2,PALPB2*
Lynch syndrome	*MLH1, PMS2, MSH2, MSH6* + 3'UTR of *EPCAM*

Gene panel	Gene sequences
Familial hypercholesterolemia	43 genes*
Retinal disorders	395 genes*
Illumina TruSight One panels	Up to 6700 genes associated with human disease in two panels

The advantages of gene panels are low costs, often better coverage—when designed well—of the genes of interest than a whole exome panel, and few variants to interpret, so that incidental findings (described below) are less troublesome.

Targeted DNA sequencing has also permitted a type of genome-wide sequencing; see the *whole exome and whole genome sequencing* subsection in the main text. *As listed in the Genomics England PanelApp at https://panelapp.genomicsengland.co.uk/panels/

Whole exome and whole genome sequencing

The Illumina TruSight One panels in Table 1 of Box 11.2 effectively represent a "*clinical exome*". Another, more long-standing, application of targeted DNA sequencing is to use biotin-streptavidin capture and sequencing of a "whole exome" from a genomic DNA sample. Such a captured exome is artificial, designed to be mostly made up of coding DNA sequences (we have more RNA genes than protein-coding genes, and protein-coding DNA accounts for only 1 % of the genome; the bias towards coding DNA is justified on the observation that pathogenic point mutations are concentrated in coding DNA). In addition to coding DNA, however, captured exomes are designed to include some untranslated sequences, notably: short intronic sequences flanking exons (to catch more splice site mutations), many (known) regulatory sequences, and sequences specifying microRNAs.

The whole exome and whole genome sequencing approaches each have their advantages and disadvantages, as listed below.

- *Whole exome sequencing* (WES) is comparatively inexpensive, but it sometimes suffers from inefficient capture; some of the desired coding sequences may be missing.
- *Whole genome sequencing* (WGS) does not suffer from the disadvantage of missing sequences, and picks up structural variants as well as point mutations. It has been used extensively in cancer studies, but it is more expensive; and with so many variants to analyze, interpretation can be much more complex unless the data are filtered.

Virtual gene panels

The curated gene panels described in Box 11.2 need to be periodically updated as new genes are added to the list or sometimes old ones are removed. An alternative approach uses *virtual gene panels*: WGS is carried out and then bioinformatic filters are applied to filter out most of the WGS dataset, retaining just the sequences of the genes of clinical interest. This approach can be expected to become more important as WGS costs fall.

Interpreting and validating sequence variants can be aided by extensive online resources

As described in the previous section, mutation scanning can be carried out at different scales. Currently, the trend is towards developing genome-wide mutation scans. Although comparatively cheap, the capture process in exome sequencing is, however, inefficient, and whole genome sequencing can be expected to replace exome sequencing once sequencing costs become sufficiently low.

Genome-wide scale comes with another cost: the number of validated variants that need to be filtered to arrive at high-probability pathogenic

variants is large, about 20 000 for exome sequencing and significantly more for whole genome sequencing (but much less for many gene panels). The process of interpreting, validating, and filtering sequence variants may consequently be very time-consuming.

To narrow down the choice of variants it may be profitable to focus on three major types of analysis: searching for precedent (by analyzing records of previously confirmed pathogenic variants), assessing sequence conservation (comparing against equivalent sequences in other organisms to draw conclusions on the functional importance of a nucleotide sequence, or often a derived amino acid that is predicted to be substituted), and determining rarity of the pathogenic variant—see **Figure 11.13A** for an overview.

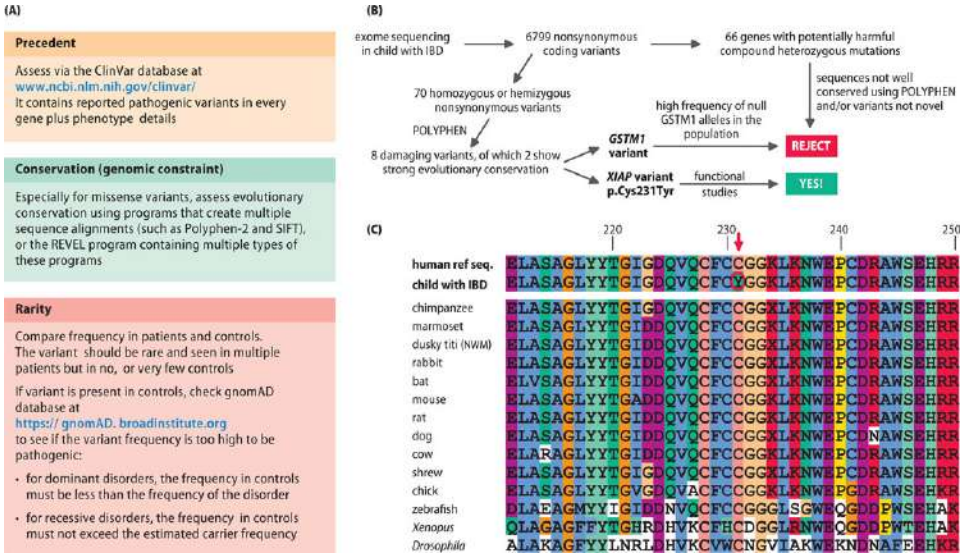


Figure 11.13 Sifting through DNA variants from a mutation scan to permit identification of a pathogenic variant. (A) Three major types of analysis. (B) An example from an early exome sequencing study in a young boy with recessive inflammatory bowel disease (IBD). POLYPHEN screening ultimately led to two strongly conserved novel missense variants. One could be excluded, because null alleles of that gene are frequent. The other, a maternally inherited variant in the X-linked inhibitor of apoptosis *(XIAP)* gene caused a p.Cys231Tyr substitution while the paternal allele carried a deletion. (C) The new *XIAP* variant appeared highly significant due to very strong evolutionary conservation of the Cys-231 amino acid, extending to

Drosophila. The mutant XIAP protein showed loss of normal function in apoptosis and *NOD2* signaling, confirming it as causative. NWM, New World monkey. (Adapted from Worthey EA et al. [2011] *Genet Med* 13:255–262; PMID 21173700. With permission from Macmillan Publishers Ltd.)

Evidence that a change is pathogenic usually comes from comparing the frequency of the candidate variant in patients and controls. Ancestrally matched control DNA samples may be used (but results are often compared with data stored in general DNA databases—see below). When sifting through variants that are also present in controls, the frequency is very important. Variants at high frequency in controls are very unlikely to be pathogenic, but the type of disorder and penetrance of mutations need to be taken into account. A sequence variant found in healthy male and female controls would usually be eliminated from consideration in a highly penetrant early-onset dominant or X-linked condition, but could contribute to disease in an autosomal recessive or a low-penetrance dominant condition.

Unlike easily identified and interpreted loss-of-function mutations, missense variants can be difficult to evaluate. A non-conservative substitution—replacing an amino acid by another of a different class—is more likely to be pathogenic than a conservative substitution. Conservative substitutions can, however, be pathogenic and nonconservative substitutions can be benign. But this is where evolutionary conservation studies can be very helpful. The concept is simple: if a sequence is functionally important there is pressure from natural selection to maintain that sequence. The sequence is subject to **genomic constraint**. Thus, if the normal amino acid is very highly conserved across a wide range of species, it is likely to be functionally very important, and a mutation producing a nonconservative amino acid change at this position becomes highly significant. Conversely, a substitution is unlikely to be pathogenic if it changes the amino acid to one that is the normal amino acid at an equivalent position in an ortholog from another species.

In Figure 11.13B,C we give an example of how conservation was particularly important in using exome sequencing to identify the genetic cause in a boy who presented aged 15 months with a life-threatening, but previously unidentified, form of inflammable bowel disease. Because of the severity of the disorder at such an early age, a recessive disorder was expected. Exome sequencing identified 16 124 DNA variants when compared against the human genome reference sequence, with a total of 6799 substitutions that were then analyzed as shown in Figure 11.13B,C to identify a pathogenic missense mutation in the *XIAP* gene. The discovery led to a change in treatment.

If a variant, especially a nonconservative missense variant, arises *de novo* in an affected individual, its candidacy as a contributor to pathogenesis is also increased, especially if it is a nonconservative missense variant. We consider segregation of variants in families below.

Clinical reporting and nomenclature of sequence variants

Diagnostic laboratories generally report DNA variants suspected to be associated with pathogenesis in five categories as follows:

- Pathogenic
- Likely pathogenic
- Uncertain significance
- Likely benign
- Benign.

In reporting nucleotide and protein variants, the HGVS nomenclature is recommended, as described in **Box 11.3**.

Standards and guidelines for the interpretation of sequence variants

The foundation for current interpretation of sequence variants was recently set by the American College of Medical Genetics and Genomics jointly with the Association for Molecular Pathology (Richards et al. [2015] under Further Reading). They proposed 16 criteria supporting pathogenicity (**P**), placing them in four groups according to the strength of the evidence thus: very strong (PVS) with one category, PVS1; strong (PS) with categories PS1 to PS4; moderate (PM) with categories PM1 to PM6; and supporting (PP) with categories PP1–PP5.

BOX 11.3 NOMENCLATURE FOR SEQUENCE VARIANTS

The nomenclature for sequence variants is described in detail in the HGVS (Human Genome Variation Society) website (http://varnomen.hgvs.org/recommendations/general). The computer program Mutalyzer (https://mutalyzer.nl) will generate the correct name of any sequence variant that a user inputs. The nomenclature for sequence variants requires a reference sequence, which may be one obtained from a database such as RefSeq, the NCBI's reference sequence database at https://www.ncbi.nlm.nih.gov/refseq. More recently, the Locus Reference Genomic (LRG) database https://www.lrg-sequence.org/ has come to be commonly used, providing stable reference sequences for reporting sequence variants with clinical implications. The nomenclature for sequence variants has three main components, as follows:

1. A reference sequence shown by a recognized accession number followed by a symbol describing the type of sequence as follows: **g**. (nuclear genome); **m**. (mitochondrial genome); **c**. (coding DNA); **n**. (non-coding DNA); **r**. (RNA); **p**. (protein)
2. A number, or a range of numbers separated by an underscore, that defines the position(s) changed within the reference sequence
3. A description of the type of change. For nucleotide sequences: a substitution (>), deletion (del), insertion (ins), duplication (dup), a deletion-insertion event (delins), or an inversion (inv). For RNA

sequence variants, nucleotides are shown in lower case. For proteins, the symbols fs, *, and ext indicate, respectively, a frameshift, termination codon, and extension of the protein sequence.

The reference sequence may have an accession number recognized by sequence databases such as NM_000249.4, the *genomic* sequence corresponding to a major *MLH1* gene transcript, or it might be a Locus Reference Genomic database reference number, such as LRG_199t1 which is the genomic sequence corresponding to a primary transcript of the dystrophin gene, the Dp247m isoform (also present in the NCBI database with the accession number NM_004006.2). An example of a full variant sequence would be: LRG_199t1:c.79_80insG (= insertion of a G between nucleotides 79 and 80 of the LRG_199t coding DNA). **Table 1** lists some examples of the nomenclature for sequence variants, but omitting the reference sequence for simplicity.

TABLE 1 SOME EXAMPLES OF NOMENCLATURE FOR DESCRIBING SEQUENCE VARIANTS

Example	Interpretation
g.19C>A	substitution of a C by an A at position 19 in the genomic sequence
c.79_80delinsTT	nucleotides 79 and 80 of the coding DNA sequence are replaced by the dinucleotide TT
c.872_875del	deletion of nucleotides 872-875 in the coding sequence
c.*57C>G	replacement of C by G at nucleotide position 57 in the 3′ untranslated region

Note that substitutions are confined to just a single nucleotide; if a CC dinucleotide at positions 79 and 80 were replaced by a TG that would be represented as 79_80 delinsTT, not as two separate substitutions at nucleotides 79 and 80.

Example	Interpretation
c.178+9A>G	replacement of A by G at the ninth nucleotide within the intron that follows nucleotide number 178 in the cDNA (the last nucleotide of the preceding exon)
c.179–3C>T	replacement of C by T at third nucleotide preceding nucleotide 179 in the cDNA (the first nucleotide of the following exon)
p.Asp107His	aspartate at amino acid position 107 is replaced by histidine
p.Gly542*	the codon specifying glycine at amino acid (a.a.) position 542 is replaced by a stop codon
p.Arg123LysfsTer34	a variant with arginine 123 as the first amino acid shifts the reading frame, replacing it with a lysine and terminating after another 33 codons

Note that substitutions are confined to just a single nucleotide; if a CC dinucleotide at positions 79 and 80 were replaced by a TG that would be represented as 79_80 delinsTT, not as two separate substitutions at nucleotides 79 and 80.

As might be expected, the strongest criterion for pathogenicity, PVS1, is evidence of a null variant—such as a nonsense or frameshifting mutation, a change to the canonical GT and AG motifs at splice sites, single- or multi-exon deletions, or a change to the initiation codon—in a gene where loss of function is a known disease mechanism. Other strong criteria for pathogenicity include PS1 and PS2. In PS1, a nucleotide change is reported to give the same missense variant as one previously reported to be pathogenic but via a different nucleotide change. An example: a **G**GA glycine codon known to undergo a pathogenic G>A mutation to give the arginine codon **A**GA, is found in another affected individual to have been replaced by a **C**GA codon, also specifying arginine. PS2 covers a *de novo*

mutation in a patient with disease but no family history, and with both paternity and maternity confirmed.

A total of 10 criteria for benign (**B**) variants were placed in three groups: stand-alone with one criterion (BA1); strong (BS) with four criteria (BS1–BS4), and supporting (BP) with five criteria (BP1–BP5). The strongest evidence for a benign variant, occurs when its frequency is exceptionally high: at >5 % in the population for criterion BA1 (as derived from global population variation data), or for BS1, at a higher frequency than could be expected for the disorder. The evidence framework for the criteria above is displayed in **Figure 11.14**.

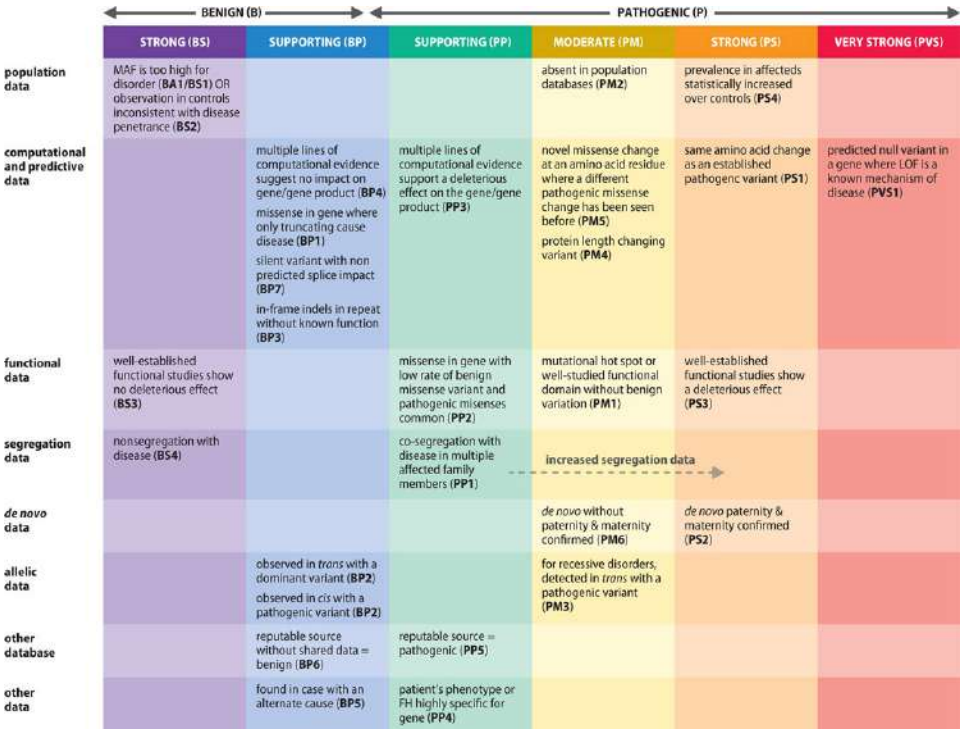


Figure 11.14 Evidence framework for criteria classifying pathogenic or benign variants. MAF, minority allele frequency. LOF, loss of function. FH, family history. Path., pathogenic. See text for description of criteria. Note that in 2020 the UK's Association for Clinical Genomic Science (ACGS) released its best practice guidelines for variant classification in rare disease; these can be accessed at https://www.acgs.uk.com/media/11631/uk-practice-guidelines-for-variant-classification-v4-01-2020.pdf. Also note that some variants—including some missense mutations, some synonymous mutations with possible splice effects, nonsense and frameshifting

mutations where the predicted termination codon lies close to the end of the coding sequence, and some intronic mutations within a few nucleotides of the intron terminating GT and AG dinucleotides—may not easily be classified and are considered variants of uncertain clinical significance. We consider such variants of unknown significance later in the text. (Adapted from Richards et al. (2015) Genet Med 17:405–424; PMID 25741868 with permission from Springer Nature.)

Note the importance of increasing amounts of segregation data in Figure 11.14, and of *de novo* mutation. For people with a dominant disorder but unaffected parents, a *de novo* mutation has a generally higher likelihood of being pathogenic than an inherited mutation, especially if paternity and maternity are confirmed. If the disorder is familial, the mutation might be checked in other family members. If the mutation does not segregate with disease, it is highly unlikely to be implicated in the disease, assuming a high penetrance. But the reverse is not necessarily true: co-segregation with disease is not evidence that a variant is pathogenic (a nonpathogenic variant at a disease locus has a 50 % chance of residing within the same allele as the true disease-causing mutation, and a 50 % chance of co-segregating with disease). Segregation with disease may need to be studied through multiple meioses in a family.

In silico resources for describing and interpreting sequence variants

We have described a few computational and database resources above, but in **Table 11.5** we summarize the range of major internet resources in this area. Note that despite the reliance on *in silico* resources, in some cases it may be necessary to validate an interpretation by laboratory analyses. Functional analyses that show a mutant protein does not carry out the normal functions can provide high confidence in pathogenicity when suitable analyses can be done, as in the example of Figure 11.13. Until recently, non-canonical splice site variants would need to be confirmed by

reverse transcriptase-PCR analyses, but powerful new programs such as SpliceAI are providing greater confidence in predicted splice variants.

TABLE 11.5 MAJOR RESOURCES FOR DESCRIBING AND INTERPRETING SEQUENCE VARIANTS ANDTHEIR CLINICAL RELEVANCE

Resource	Description
ClinGen	At www.clinicalgenome.org. An aid to exploring the clinical relevance of genes and variants. Includes assessment of gene-disease validity; evaluation of gene dosage sensitivity, with the HI haploinsufficiency scores predictive model, based on functional, evolutionary and network properties; clinical actionability; curated variants. Reviewed at PMID 26014595.
ClinVar	At www.clinvar.com. Gives reports of the relationships between human variants and phenotypes, with supporting evidence. Reviewed at PMID 26582918 and 31777943.
gnomAD	At https://gnomad.broadinstitute.org/. The Genome Aggregation Database (gnomAD), the major resource on human genetic variation, was constructed by aggregating international exome and genome sequence data. It has a gene/missense constraint track with pLI and Z-scores (a measure of constraint at the gene level; for a missense variant, a Z score >3.09 is deemed significant, indicating that it is intolerant to variation).The pLoF program measures a transcript's intolerance to loss-of-function (LoF) variation, with observed/expected (oe) values presented, low values indicating constraint against variation.
HGMD	At http://www.hgmd.cf.ac.uk/ac/index.php.The Human Gene Mutation Database collates published gene lesions responsible for human inherited disease. Available as a free version and as a professional version requiring a subscription. Reviewed at PMID 32596782.

Resource	Description
Mastermind Professional	At https://www.genomenon.com/mastermind/. Powerful publication resource covering genetic variations (and disease-gene associations). Enables graphical and text querying to scan the literature for pathogenic variants, and assists the sensitivity and reproducibility of clinical variant interpretation. Requires a subscription. Reviewed at PMID 33281875.
REVEL	Freely available at https://sites.google.com/site/revelgenomics. Powerful software that predicts pathogenicity of missense variants based on combining scores from PolyPhen-2, SIFR, PROVEAN and 10 other computational tools. Reviewed at PMID 27666373.
Alamut Visual Plus™	At https://www.interactive-biosoftware.com/alamut-visual-plus/. Proprietary software that curates data from multiple sources (ClinVar, dbSNP, COSMIC, Mastermind, PubMed) plus offers both high-quality missense predictors and splicing predictors (including SpliceSiteFinder-like, MaxEntScan and others) in one place.
SpliceAI	At https://github.com/Illumina/SpliceAI. This deep neural network method can efficiently model mRNA splicing from a genomic sequence and can accurately predict noncoding cryptic splice site mutations in patients with rare genetic diseases—see PMID 30661751.
Coding Constrained Region	Measure of regional missense constraint derived from gnomAD data. Data available at https://github.com. See PMID 30531870 for the map of constrained coding regions (CCR) in the human genome.

Incidental findings and variants of uncertain clinical significance

Genome-wide mutation scanning—whether at clinical exome, whole exome or whole genome levels—is especially prone to some problems that can provide clinical and/or ethical difficulties. One is *incidental findings*, where medically important findings are made that are unrelated to the medical reason for which the genetic test was ordered. Mutation scanning in a patient with a heart defect may disclose, for example, a harmful mutation in a hereditary cancer gene. We consider this when we look at ethical issues related to genetic testing in Section 11.5.

Another issue concerns how to manage *variants of uncertain clinical significance* (VUS) that become a real issue in genome-wide sequencing. Each one of us will have a large number of VUS—see the legend to Figure 11.14 for some examples of common classes of VUS). While most VUS might be expected to be benign, reporting them can cause great anxiety to the patient, leaving both patient and referring physician with unanswerable questions. If they do not get reported, however, the chance of revisiting them at a later stage, when more is known, goes away. That can be undesirable: should there be a pathogenic VUS, responsive action might have been able to be taken to prevent or ameliorate future clinical symptoms.

Detecting aberrant DNA methylation profiles associated with disease

Aberrant epigenetic changes are heavily involved in the development of cancers. As detailed in Section 6.3), they also make important contributions to several inherited disorders. Often, they occur in response to some genetic change. The inherited disorders that show aberrant DNA methylation include notably imprinting disorders (according to the sex of the transmitting parent, alleles at imprinted gene loci are subject to epigenetic silencing and hypermethylation). In the case of disorders such as Prader-

Willi syndrome and Angelman syndrome, for example, testing for aberrant DNA methylation is the front-line molecular genetics test to confirm the diagnosis. That is so because essentially all affected individuals show aberrant cytosine methylation profiles for the relevant chromosome region (15q11-q12 in both cases), whether there has been a microdeletion on the normal chromosome or uniparental disomy for that chromosome.

Different methods can be used to detect aberrant DNA methylation. Bisulfite PCR methods are inexpensive and often used as a quick way to exclude negative cases for Prader-Willi and Angelman syndromes. MS-MLPA, a methylation-sensitive variant of the MLPA method that was described in Section 11.2, allows simultaneous semi-quantitative detection of the methylation status of genes and their copy number.

Bisulfite sequencing

Methylated and unmethylated cytosines can be distinguished by making the DNA single stranded and treating it with sodium bisulfite (Na_2SO_3). Under controlled conditions, the unmethylated cytosines are deaminated to produce uracils, but 5-methylcytosines remain unchanged. After treatment with sodium bisulfite, the relevant region can be amplified by PCR, during which newly created uracils are read and propagated as thymines. New DNA strands are synthesized without incorporating methyl groups so that any retained methylated cytosines in the template DNA are propagated as unmethylated cytosines. That allows different ways of distinguishing the methylated cytosines from the original unmethylated cytosines.

Figure 11.15 shows how, after treatment with sodium bisulfite, samples can be amplified by PCR and sequenced to distinguish methylated cytosines from unmethylated cytosines. Methylation-specific PCR assays can also be devised by designing alternative PCR primers to have 3¢ nucleotides that are specific for one of the variable nucleotides after treatment with sodium bisulfite (that is, a U or T versus a C).

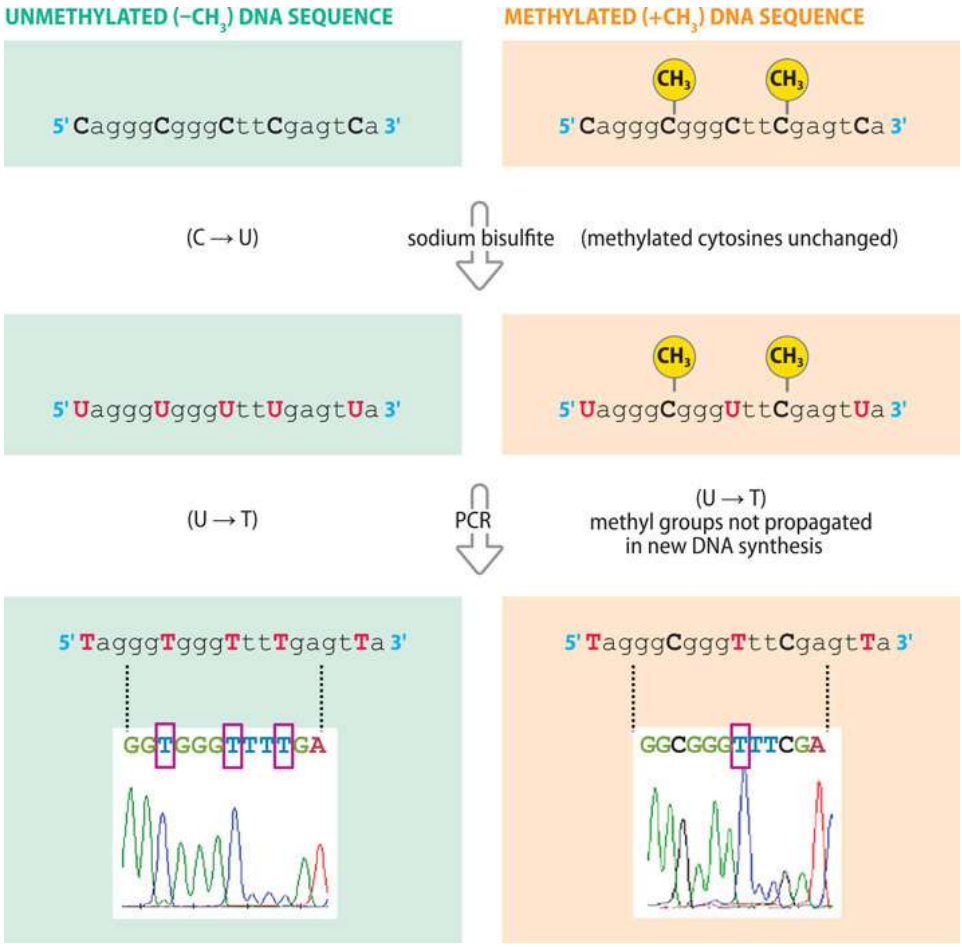


Figure 11.15 Distinguishing methylated cytosines from unmethylated cytosines with the use of sodium bisulfite. Sodium bisulfite converts unmethylated cytosines (left panel) to uracils; after DNA replication in a PCR reaction, they become thymines in newly synthesized DNA. Sodium bisulfite does not react with methylated cytosines, which remain unchanged (right panel). Newly synthesized DNA strands in a PCR reaction are not methylated, and so although the starting DNA is methylated in the right panel, the PCR product is unmethylated. DNA sequencing can identify all unmethylated cytosines because after treatment with sodium bisulfite each unmethylated C becomes a T (shown by boxes in sequencing panels at the bottom); if the cytosines are methylated, the sequence obtained is the same as in DNA that has not been treated with sodium bisulfite. DNA sequencing of PCR products is therefore one way of distinguishing between the two patterns. Alternative assays use methylationspecific PCR by designing primers with a nucleotide at the 3¢ end that corresponds to a variable site, pairing with either U/T (from an unmethylated cytosine that has been chemically converted by

sodium bisulfite) or C (representing an originally methylated cytosine). Other assays take advantage of methylation-sensitive restriction enzymes (see Text).

Note that very sensitive detection of DNA methylation status can be achieved by bisulfite treatment of sample DNA followed by the *pyrosequencing* method described above, and is often used in cancer studies.

Methylation-sensitive MLPA (MS-MLPA)

This is the front-line test for confirming doubtful cases of Angelman and Prader-Willi syndromes (especially rare mosaics), assessing DNA methylation while simultaneously being able to pick up associated microdeletions. It is also used to detect imprinting center microdeletions, and is useful in cancer studies for rapid assessment of promoter hypermethylation.

MS-MLPA is a slight modification of the MLPA method (described previously in Figure 11.5) that makes use of the methylation-sensitive restriction nuclease *Hha*I. The left MS-MLPA probes are designed to contain a *Hha*I recognition sequence, GCGC, and when denatured and hybridized to a desired target sequence containing clustered CpG dinucleotides, a heteroduplex of MLPA probe and desired target sequence has a recognition sequence for *Hha*I. If the target sequence is methylated, the *Hha*I enzyme cannot cleave the recognition sequence and amplification occurs as in normal MLPA. If it is not methylated, *Hha*I cleaves the heteroduplex and amplification cannot occur (see **Figure 11.16**).

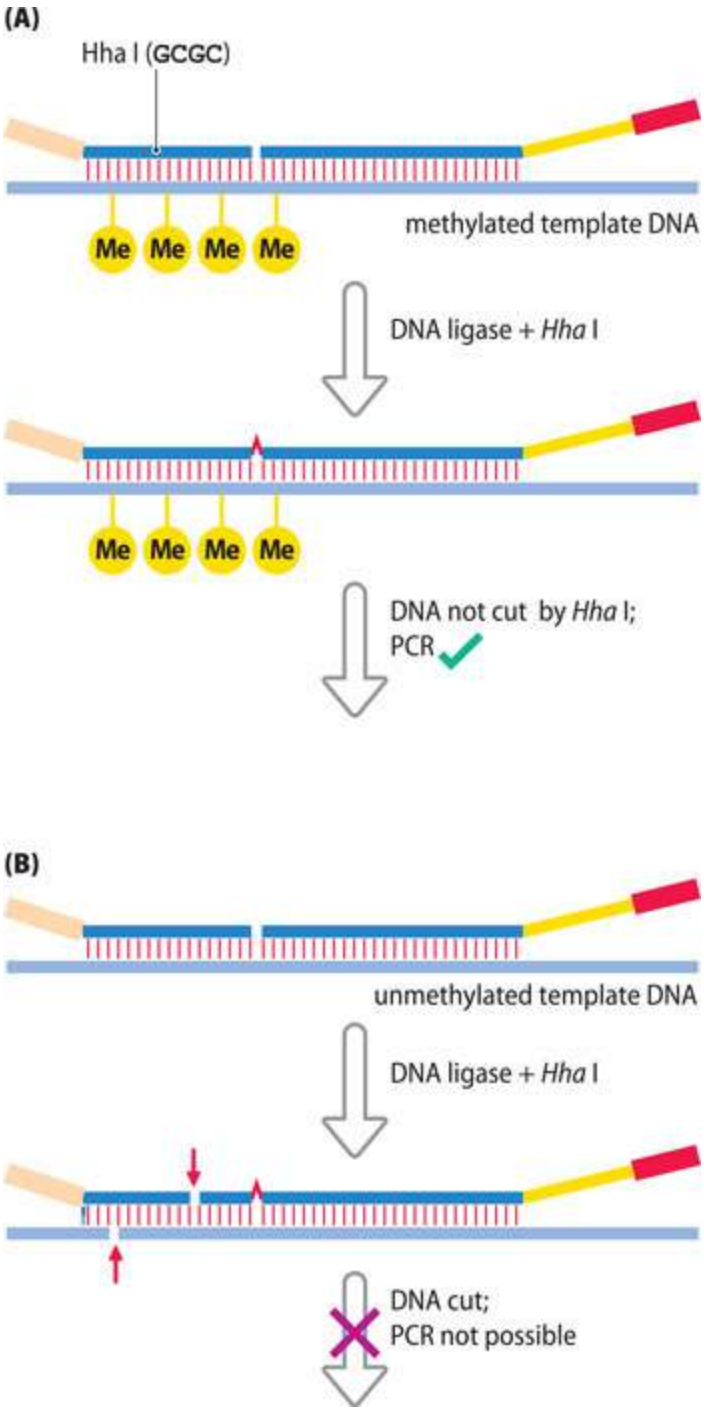


Figure 11.16 The basis of methylation-sensitive MLPA. In this modification of the MLPA method (see Figure 11.5), a *Hha*I restriction site has been engineered into the left MLPA probe. The two MLPA probes are designed to hybridize to a DNA methylation-prone GC-rich sequence on the template DNA, including a CpG within the GCGC sequence complementary to the GCGC of the *Hha*I recognition sequence. If that CpG is

methylated, *Hha*I cannot cut the DNA, and PCR amplification occurs as normal. If it is unmethylated, *Hha*I cuts the DNA and PCR is not possible.

After hybridization the MS-MLPA reaction is split into two parts: one is treated as a normal MLPA reaction to assess copy number; the other treated with *Hha*I to assess methylation status.

11.4 GENETIC AND GENOMIC TESTING: ORGANIZATION OF SERVICES AND PRACTICAL APPLICATIONS

In the previous two sections we covered the technology of genetic and genomic testing, describing how chromosome abnormalities and large-scale copy number and structural variation are detected in Section 11.2, and how testing for point mutations and DNA methylation are carried out in Section 11.3. In those sections we briefly alluded to different clinical settings in which those testing methods are deployed, and the different scales of testing, from individuals and families to communities and populations. Here, we now take a close look at the clinical context of genetic and genomic testing, and the organization and development of these services. We begin with a radical change that is transforming some genetic services into genomic medicine services. Thereafter we focus on different levels at which testing is offered.

The developing transformation of genetic services into mainstream genomic medicine

Genetic services evolved in many countries in the early 1960s to translate new chromosome findings into clinical services. Initially, the services were staffed by cytogeneticists and clinicians; DNA testing was added in the mid-1980s. In some countries services developed around particular genetic conditions, such as the thalassemias.

The 1990s saw development of cancer genetic services. Testing for predisposition to breast, bowel, and ovarian cancers served to guide enhanced screening recommendations. A decade later, similar developments took place in cardiogenetic services, as heritable causes of sudden cardiac death came to be explored, including arrhythmias and cardiomyopathies. Diagnostic approaches to the child with developmental delay, possibly with dysmorphic features, have become ever more granular since initial cytogenetic approaches. Of course, the developments over the last decade took place at the same time as our understanding of the molecular basis of genetic disorders increased exponentially. Parallel exponential improvement in available technologies was achieved in terms of both speed and cost.

About 20 years ago, at the turn of the new millennium, genetic testing was available for few disorders; patients were generally seen by clinical geneticists before having a test. More recently, advances in sequencing technology and the development of whole exome and, genome sequencing, have enabled broad genetic testing on an individual patient basis within a clinically useful time-frame. Initially, implementation of such tests was carried out by way of clinical research studies (such as the UK's Deciphering Developmental Disorders project, for example). More recently exome sequencing has been utilized as a clinical diagnostic test.

In some countries, genome sequencing is transitioning to routine healthcare. The UK has been in the vanguard of these developments, but other countries have also invested substantially in establishing national genomic medicine initiatives to address implementation barriers and transition testing from centers of excellence to mainstream medical practice (see **Box 11.4**).

BOX 11.4 NATIONAL GENOMIC MEDICINE INITIATIVES AND THE DEVELOPMENT OF MAINSTREAM GENOMIC MEDICINE SERVICES

Within the last decade the governments of at least 14 countries have invested billions of dollars in establishing national genomic medicine

initiatives to address barriers to implementation and transition testing from centers of excellence to mainstream medical practice. In countries such as the UK, France, Australia, Saudi Arabia, and Turkey, necessary resources have been developed enabling genome sequencing of large numbers of patients with rare diseases and cancer. Other countries, such as the US, Estonia, Denmark, Japan, and Qatar, have invested in population-based sequencing projects with return of results to participants.

In the UK, the momentum towards a national genomic medicine service began in 2009 when the House of Lords Science and Technology Committee issued a report calling for a strategic vision for developing genomic medicine. Follow-up workshops involving major stakeholders led to two further reports on *Genetics and Mainstream Medicine* in 2011 and *Genomics in Medicine* in 2012 that set out a route to effective integration of genomics across the NHS. An NHS Long Term Plan then laid out a vision to enable it to harness the power of genomic technology to improve the health of the population and to be the first national healthcare system to offer whole genome sequencing as part of routine care.

In 2013 the UK Government established Genomics England with a mandate to sequence 100 000 genomes from patients with a range of rare diseases and seven common cancers. The objective of sequencing 100 000 whole genome sequences was met in December 2018 (but the communication of results to patients is still ongoing at the time of writing). The follow-up vision now is to sequence half a million genomes by 2023/24 with the aim of improving healthcare for rare diseases and cancers as well as other specific aims such as developing the UK's newborn screening program, developing a pharmacogenomic service and the early detection and treatment of high-risk conditions such as familial hypocholesterolemia, and to link seamlessly with research into genomics so that findings in this setting can be rapidly incorporated into healthcare.

Having previously only been available to citizens via initiatives such as the 100 000 Genomes Project, genomic sequencing is now being offered

by the NHS's Genomic Medicine Service as a standard test—see **Figure 1** for the information about testing offered to patients.

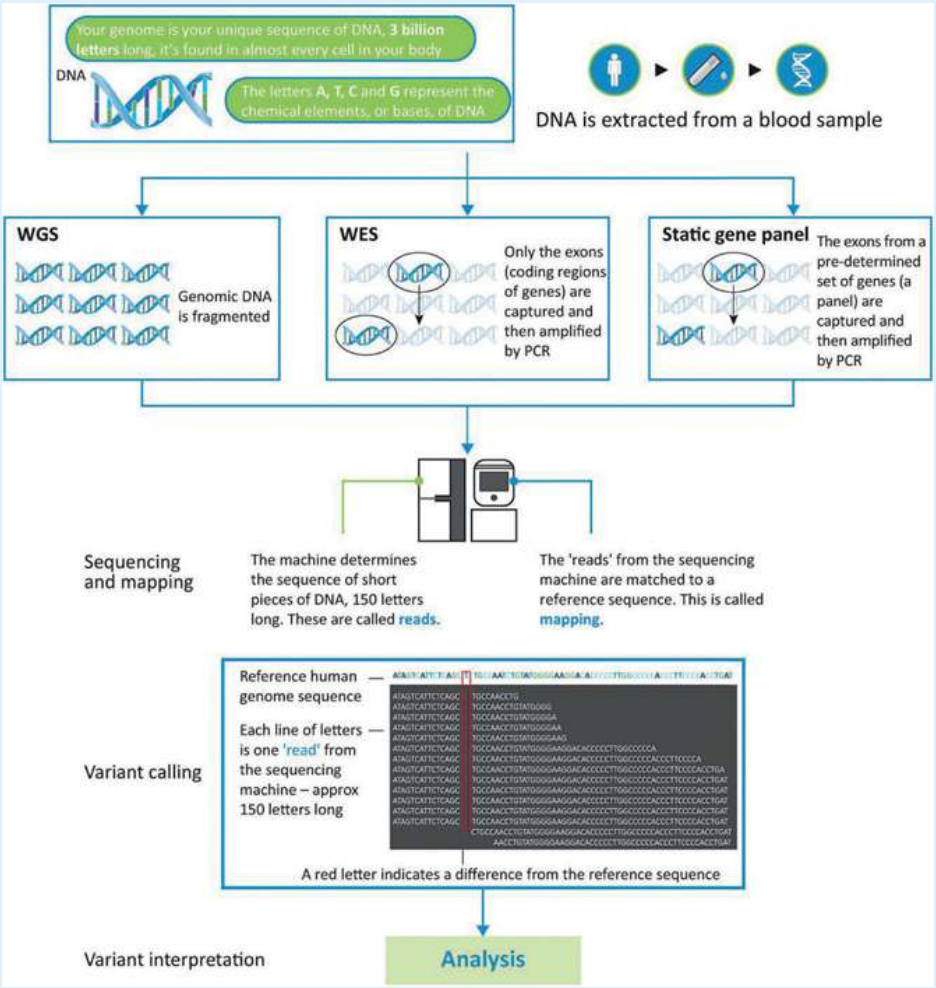


Figure 1 Outline of genomic testing approach offered to NHS patients. A national genomic test directory in England (https://www.england.nhs.uk/publication/national-genomic-test-directories/) specifies which tests, ranging from single gene to whole genomes, are commissioned for use (© genomics England).

A national genomic medicine service embedded throughout mainstream specialties has the potential to enable quicker diagnoses (reducing the length of the diagnostic odyssey for many children with rare diseases, to match people to the most effective medications and interventions (reducing the chances of an adverse drug reaction), and to increase the survival from cancer through earlier diagnoses and more targeted

therapies. This is no small undertaking. Setting up a consistent and equitable national genomic medicine service is a laudable aim, but one that will require a close focus on implementation issues; attempts to do so in the UK during a global pandemic have, almost inevitably, led to delays to date.

Genetic and genomic testing are being offered now by a much wider range of healthcare practitioners. And an agenda of *mainstreaming genetics* envisages that any branch of healthcare practice should be sufficiently versed in the implications of particular genetic tests to offer such testing directly to their patients. Increasingly, therefore, diagnostic genetic testing will become the responsibility of the clinicians to whom patients are initially referred, and clinical genetic services will take on a molecular pathology role—being involved in multidisciplinary discussions about a patient perhaps, or organizing cascade predictive testing of family members where relevant.

As a result of mainstreaming genetic services, clinicians most skilled in particular groups of disorders will be able to add genetic investigation to the diagnostic tests available to them to personalize their treatments to a greater degree. For example, the discovery that PARP-1 (poly[ADP-ribose] polymerase) inhibitors are particularly effective in treating breast cancer in patients with *BRCA1* and *BRCA2* mutations allows this particular treatment to be initiated at an earlier stage if *BRCA1* and *BRCA2* testing is available promptly at breast cancer diagnosis. In cells with either of these mutations, homologous recombination (one of the two major DNA repair methods) is nonfunctional, but base-excision repair is unaffected. PARP-1 inhibition disables base-excision repair, and thus cells with *BRCA1* or *BRCA2* mutations are no longer able to repair DNA. In the past, tests for *BRCA1* or *BRCA2* mutations were often undertaken largely so that relatives might be offered predictive testing, but the examples above illustrate how rapid genetic testing will increasingly be part of the mainstream specialist

approach to deciding optimal treatment that is personalized to the particular patient's genetic make-up.

An overview of diagnostic and pre-symptomatic or predictive genetic testing

In rare, single-gene conditions, genetic tests are often directed at persons presenting with a clinical problem for *diagnostic testing*. Once the causative (or strongly predisposing) genetic variant has been found, close relatives of that person may be offered some type of predictive genetic testing to see whether or not they have inherited the variant in question. This in turn provides information on whether action is required, and what steps might need to be taken to reduce the likelihood of the clinical problem arising, or to detect it in an early, more treatable stage. Depending on the penetrance, such testing is often described in two slightly different ways:

- *Pre-symptomatic testing* is used when a patient possessing the mutant allele(s) will at some stage develop the condition in question, such as in the case of Huntington disease.
- *Predictive testing* is used if patients are at high *risk* of developing symptoms, such as in the case of carriers of pathogenic *BRCA1* or *BRCA2* variants.

The distinction between diagnostic and predictive testing is relied upon in policy to avoid genetic discrimination (as set out, for example, in the UK government's code on genetic testing and insurance at https://www.abi.org.uk/globalassets/files/publications/public/genetics/code-on-genetic-testing-and-insurance-final.pdf). But that same distinction is becoming less clear as whole genome approaches increasingly unearth both diagnostic and predictive findings. Furthermore, the latter will often be significantly less certain in their predictions (with reduced penetrance for example) than the single-gene genetic tests from which services have amassed the most experience to date.

For some diseases the clinical benefit of predictive testing is clear. Take, for example, familial hypercholesterolemia (OMIM 143890), an autosomal dominant disorder commonly caused by pathogenic mutations in the *LDLR* (low-density lipoprotein receptor) gene. Affected individuals normally develop premature cardiovascular disease in the third decade, but early detection of a pathogenic *LDLR* mutation offers the possibility of prevention by lowering LDL-cholesterol through dietary changes and medication. That has led to recommendations of *cascade testing* (testing of relatives after a genetic condition has been identified in a family), either by measuring LDL-cholesterol, or by testing for the familial *LDLR* mutation from as young as the age of 10 years.

Other examples of the balance towards benefits over risks include colorectal cancer syndromes such as Lynch syndrome and familial polyposis coli, which are dominantly inherited. Early detection of a germline mutation that predis-poses to these diseases can be followed up by regular colonoscopy surveillance. By identifying and surgically removing polyps before they grow and become dysplastic, the risk of developing late-stage cancer is much reduced. There is a small risk of perforation of the bowel during the required colonoscopy, which needs to be factored into this balance, but this risk is generally lower in the young population who need to be screened before national bowel screening programs kick in, and be in the hands of experienced colonoscopists who run dedicated family history screening programs.

In other familial cancer syndromes, the benefits of predictive testing may be less clear. For example, Li Fraumeni syndrome due to germline *TP53* variants will detect those at high lifetime risks of various cancers including breast and sarcomas. However, the evidence that screening (for example, whole body MRI screening) will make any cancers more amenable to treatment is still quite limited.

Note that the surveillance carried out for some familial cancer syndromes (see **Clinical Box 16**) does not prevent cancer—its aim is to identify early cancers while they are still amenable to therapy. To reduce their risk of developing cancer, women from families with a known *BRCA1* or *BRCA2*

variant associated with risk may opt for bilateral mastectomy and/or surgical removal of the ovaries together with the associated fallopian tubes. Predictive testing would be indicated before making this decision, and indeed is not usually offered without a positive test result for the familial mutation.

CLINICAL BOX 16 LYNCH SYNDROME AND FAMILIAL (*BRCA1/BRCA2/PALB2*) BREAST CANCER: CANCER RISKS AND CANCER SCREENING

LYNCH SYNDROME (HEREDITARY NONPOLYPOSIS CANCER)

The diagnosis is determined on the basis of the pattern of cancers in a family and the age at diagnosis (at least one diagnosis under the age of 50 years), or on the finding of microsatellite instability in tumor tissue, or on immunohistochemistry evidence of abnormal gene expression. Although colorectal cancer is the commonest cancer in the condition, there are several associated cancers (Table 1).

TABLE 1 CANCER RISKS IN LYNCH SYNDROME COMPARED WITH THE NORMAL POPULATION

CANCER	GENERAL POPULATION RISK (%)	RISK IN LYNCH SYNDROME (%)
Colorectal	5.5	20-80
Endometrial	2.7	20-60
Gastric	<1	1-10
Ovarian	1.6	9-15
Hepatobiliary tract	<1	2-7
Urinary tract	<1	4-5
Small bowel	<1	1-4

CNS, central nervous system.

CANCER	GENERAL POPULATION RISK (%)	RISK IN LYNCH SYNDROME (%)
Brain and CNS	<1	1-3

CNS, central nervous system.

As testing has become more widespread and initiated with weaker family histories, the risk profiles in Lynch syndrome have also widened. This is so because by selecting only the strongest family histories for genetic testing, other familial factors affecting cancer incidence were selected in the process. Screening a general population—unselected for family history—would mean that the associated cancer risks would on average be lower. As yet, there is no widespread population screening for such dominant cancer genes, but the concept is of importance, since many people advocate that their incidental discovery during genome-wide scans should be disclosed [ACMG53/57/73—https://www.ncbi.nlm.nih.gov/refseq] or that they be specifically sought as "additional" findings during genomic approaches in the diagnosis of other conditions.

A typical screening protocol for affected and high-risk individuals would be colonoscopy every two years, starting at the age of 25 years, followed by additional endoscopy examination of the esophagus, stomach, and duodenum from 50 years of age. The efficacy of screening for other associated tumors is not proven. Women who have the condition, or who are at high risk, may opt for total hysterectomy and surgical removal of ovaries plus fallopian tubes after completion of their family.

FAMILIAL BREAST CANCER DUE TO *BRCA1/BRCA2/PALB2* VARIANTS

In the general UK population, the lifetime risks of breast cancer and ovarian cancer are roughly 12 % and 2 %, respectively. A person aged 20–25 years with a pathogenic *BRCA1, BRCA2*, or *PALB2* variant has a

roughly 70 % risk of going on to develop cancer, notably breast or ovarian cancer (Table 2).

TABLE 2 CANCER RISKS FOR CARRIERS OF PATHOGENIC BRCA1 /BRCA2/ PALB2 VARIANTS

Cancer type	Lifetime (to age 80 years) risk (%)		
	BRCA1	***BRCA2***	***PALB2***
UNAFFECTED CARRIERS			
Breast cancer	60–90	30–85	40–60
Ovarian cancer*	30–60	10–30	2–10
Male breast cancer	0.1–1	5	0.2–6
Prostate cancer	8**	25	7
Other cancers	<5	<5	<5
AFFECTED WOMEN CARRIERS (WITH UNILATERAL BREAST CANCER)			
Cancer in other breast	50% (overall 5-year risk =10%	50% (overall 5-year risk = 5–10%	7

* Majority of lifetime risk after 40 years of age.

** Similar to population risk.

When a pedigree indicates a high likelihood of familial breast or ovarian cancer, a typical screening program would commence with annual mammography at 30–40 years of age in women at high risk. Mammography is less sensitive in women aged less than 40 years and so would often be accompanied by MRI breast screening from 30 years of age or earlier. Bilateral mastectomy reduces the risk of developing breast cancer by 95 %, but cancer can still occur in remaining breast tissue on the chest wall.

There is very limited evidence that ovarian screening allows the detection of cancers at a more treatable stage. Women may consider the surgical removal of ovaries and fallopian tubes—it reduces the risk of ovarian cancer by 95 % (with a small residual risk of primary peritoneal carcinoma) and depending on the age at which it is done, it may also reduce the risk of breast cancer. Ovarian cancer due to these genes is rare below the age of 40, and the risk of bilateral salpingo-oophorectomy at younger ages is significant because of the premature surgical menopause this precipitates.

Cascade testing

Cascade testing means testing of relatives after the identification of a genetic condition in a family. The relatives might be at risk of going on to develop the same single-gene disorder (predictive testing, see above). Unaffected relatives may also be at risk of transmitting a disorder if they carry a harmful allele (heterozygote carriers in recessive disorders, nonpenetrance in dominant disorders) or a balanced translocation.

Different issues need to be considered. How important is it for the relatives to be made aware of the information on the basis of the severity of the condition and the level of risk of a relative developing the condition, or having a child with the condition? What treatment or intervention is available to those who have inherited the factor in question? How might the information change things? How easy will it be for family members to pass information on to relatives? And can, and should, health professionals be involved in such communication?

Take the example of a child with multiple malformations and developmental delay who has inherited unbalanced chromosome translocation products from a parent with a balanced translocation. The translocation will be explained to the parents along with information about their future pregnancies, and also the possibility of other family members

carrying the same balanced translocation. In addition to addressing questions from the couple about the risk to future pregnancies and about their child's future, health professionals need to consider which additional family members should be contacted who might have the same balanced translocation (and be at risk of producing children with unbalanced translocation products), and how to go about this. The same principles apply to cascade testing for carriers of autosomal or X-linked recessive disorders.

Predictive genetic testing in children

Weighing the relative merits and disadvantages of having a predictive test for an adult-onset condition is not easy. Individuals seeking such testing many years before interventions—such as mammography—could be offered, would often, on reflection, delay the predictive test until just before the time it would impact screening recommendations. Alternatively, refining reproductive risks might be a reason to seek predictive testing in conditions where the evidence for interventions is more limited. Predictive testing of children is appropriate when the onset of the condition is usually in childhood, as in the case of multiple endocrine neoplasia (where screening recommendations start from the age of 5). But for later onset conditions it is important to have a conversation with the parents to plan the optimum timing of a test.

For tests, where a significant proportion of adults go on to make a considered decision not to have the test, such as predictive testing for Huntington disease, it is important to preserve the decision until the child is competent enough to do so themselves. A plethora of international guidelines recommend that predictive tests for adult-onset disorders should not be undertaken in children unless a medical intervention applicable to children is both possible and shows clear medical benefit (as in familial hypercholesterolemia). One example of such guidance—from the British Society of Genetic Medicine and currently under revision—gives worked

case examples and practical suggestions for consultations with parents requesting such testing (https://www.bsgm.org.uk/about/our-history/).

Pre-symptomatic genetic testing for conditions whose course cannot be altered by medical intervention

There are no current interventions to delay the onset of most of the late-onset neurological disorders. Nevertheless, Huntington disease, a devastating neurological disorder that often does not manifest itself until later stages in life, was one of the first conditions for which predictive testing was offered. After initial concerns that individuals testing positive for this disorder might take their own lives, predictive testing was introduced with caution. Several sessions are usually scheduled with a genetic service to explore the pros and cons of such testing; in these settings experience has shown that predictive testing precipitates a catastrophic event—suicide, suicide attempt, or psychiatric hospitalization—in less than 1 % of cases.

The uptake of testing in people at 50 % risk of Huntington disease is about 10–20 %, interestingly much lower than the uptake imagined prior to the test being available. Some evidence also suggests that initial thoughts about testing are more enthusiastic than considered decisions taken after a discussion about the pros and cons. Young adults who undergo testing generally do so to assist in making career and family choices. Another group opting for testing are those who have reached the age by which signs and symptoms would usually have presented; they wish to be tested so that they can reassure their children and grandchildren that the condition has not been passed down their branch of a family.

Whilst pre-symptomatic testing for highly penetrant conditions such as Lynch syndrome or Huntington disease is often pitched as a “yes” or “no” result, for any condition with a penetrance of <100 %, individuals who test positive for the gene variant may never develop the condition. And as population genetic testing becomes more widespread (and is not selected on

the basis of pheno-type), lower penetrance variants of these conditions will be found, and communication around risk, and the interventions available to manage it, is becoming more complex.

The different ways in which diagnosis of genetic conditions is carried out in the prenatal period

Couples who have a family history of a serious genetic disorder usually want to know whether they are at risk of having an affected child. If they are at risk, they might choose not to have children at all, or to adopt an unrelated child. Other times the genetic condition in question can be avoided through the use of egg or sperm donation, or by preimplantation diagnosis so that only healthy embryos are selected, as described below. Since the latter often involves the financial burden of private fertility services, some will choose natural conception and prenatal diagnosis in which the fetus is tested to see if it has inherited the genetic condition. Given that the opportunities for therapies *in utero* remain very limited, such prenatal diagnosis is usually offered on the understanding that a pregnancy will be terminated if the fetus is affected, though clearly this will always ultimately be a pregnant woman's decision.

Careful risk assessment and communication of the options is key in prenatal diagnosis (Box 11.5). Accurate predictive genetic testing is possible for single-gene disorders in which the major genetic variant contributing to disease has been identified in an affected family member. Prenatal diagnosis may also be carried out in situations in which there is an increased risk of transmitting a chromosomal aneuploidy (advanced maternal age is an important risk factor). Or one parent might have been identified as a carrier of a balanced translocation, and there is a risk that a fetus with unbalanced translocation products might be viable but have severe problems.

Traditionally, prenatal diagnosis has involved collecting a sample of fetal tissue recovered by an invasive procedure. A sample may be taken from the chorion (the outermost extra-embryonic membrane), and fetal DNA can be

isolated from the cells obtained (**Figure 11.17A**); there is a roughly 1 % excess risk of miscarriage. The sample can be taken any time in the pregnancy from 11 weeks onward, but typically in the first trimester (to allow the possibility of early termination of pregnancy).

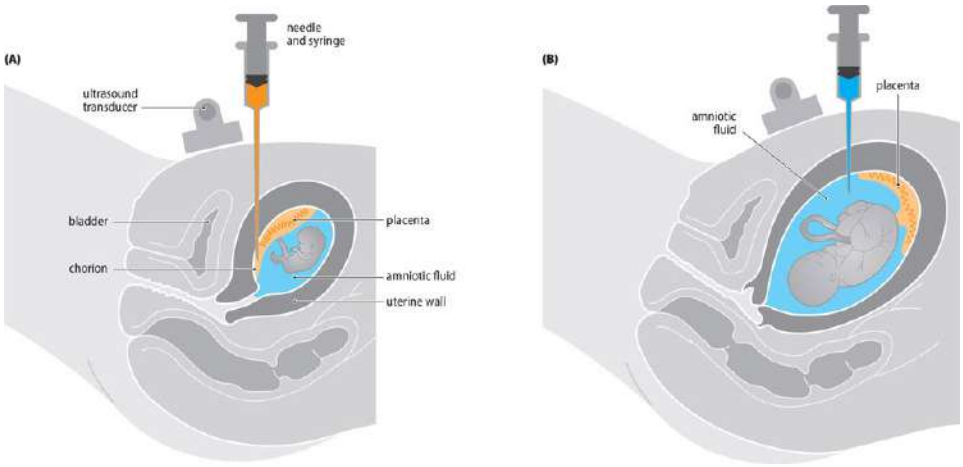


Figure 11.17 Invasive prenatal diagnosis using chorionic villus sampling or amniocentesis. (A) Chorionic villus sampling. As shown here, this is usually carried out by a transabdominal approach guided by ultrasound under local anesthetic. (B) Amniocentesis.

Amniocentesis is the other major alternative sampling method, and it also has a small risk of miscarriage. A sample of amniotic fluid is taken at, or close to, 16 weeks of gestation (Figure 11.17B); it provides fetal cells that are processed to give either chromosome preparations to check for chromosome abnormalities, or fetal DNA samples for analysis.

Pre-implantation genetic testing can also be carried out to prevent the transmission of a harmful genetic defect by using *in vitro* fertilization (IVF). We describe that separately below.

CLINICAL BOX 17 GENETIC CONSULTATIONS AND GENETIC COUNSELING

Genetic consultations

Genetic consultations often start with a child or adult with a genetic disorder—at other times consultations are initiated by a person's concern about other family members. Rather than focus on his or her own health issues, the person may bring up the subject of a family history of a medical problem such as cancer. Together, the patient and clinician construct the pedigree, the basic tool in the genetics clinic. Diagnoses are then confirmed by using, for example, cancer registries, or death certificates in the case of deceased individuals, or by requesting consent to access medical information of living relatives.

Confirming family history diagnoses can be really important in risk assessment. For example, a purported family history of bowel cancer may in fact be diffuse gastric cancer, or ovarian cancer. In that case different genes should be scrutinized, and the geneticist should consider whether, and which, genetic tests are appropriate, and whom it is most appropriate to test first. If there is a relative who has the disorder, it would often be more appropriate to test the relative first to establish the causative mutation, which would then be the basis of a predictive test.

Genetic counseling

Parents at risk of having a child with a genetic disorder, and affected individuals and relatives of an affected family member, may benefit from **genetic counseling**, the process by which they are informed of the consequences and nature of the disorder, the probability of developing or transmitting it, and the options open to them. Genetic counseling may be provided by doctors or by professionals specifically trained as genetic counselors who may have a science or nursing background.

The counselor aims to provide the necessary information to help family members to make a decision based on a patient's values and circumstances, rather than direct them toward a particular decision. Such non-directive approaches have been important in difficult decisions about termination of pregnancy, or predictive testing for untreatable conditions. Recently, however, the evidence basis of certain interventions based on

genetic testing has been improving (for example, regular surveillance improves the life expectancy of people with familial polyposis coli), and so a certain directiveness on the part of the counselor may be more appropriate.

As well as offering general support, the counseling process has at its core the determination of risks of a condition. That may be relatively simple for single-gene disorders based on Mendelian principles, but as detailed in Section 5.3 there are often complications, such as lack of penetrance or variable expressivity.

The risk estimate may be determined by a Bayesian calculation in which a prior probability (such as the risk predicted from Mendelian principles alone, for a single-gene disorder) is modified by some other relevant information. For an X-linked recessive disorder, for example, the daughter of an obligate carrier would have a 50 % risk of herself being a carrier. However, the carrier risk for a woman whose maternal grandmother is an obligate carrier but whose own mother's status is unknown (a 50 % chance of being a carrier) can be modified by circumstance. Such Bayesian calculations are now often computerized (for example, see Cambridge University's CanRisk program at https://canrisk.org/). But an understanding of the basic principles remains key to sense-checking the outputs of such programs in case of incomplete/inaccurate data entry.

In **Figure 1**, individual I-2 is an obligate carrier of the X-linked recessive condition because she has two affected boys. III-3 is concerned that her mother, II-3, might be a carrier. Because we do not know her status, II-3 has a 50 % chance of being a carrier; if she were a carrier, she would have a 50 % chance of transmitting the mutant allele to III-3. That is, the probability that III-3 is a carrier, based on this information alone, would be 50 % × 50 % = 25 %. On drawing the pedigree, however, III-3 is found to have four brothers, none of whom are affected—this additional conditional information alters the risk.

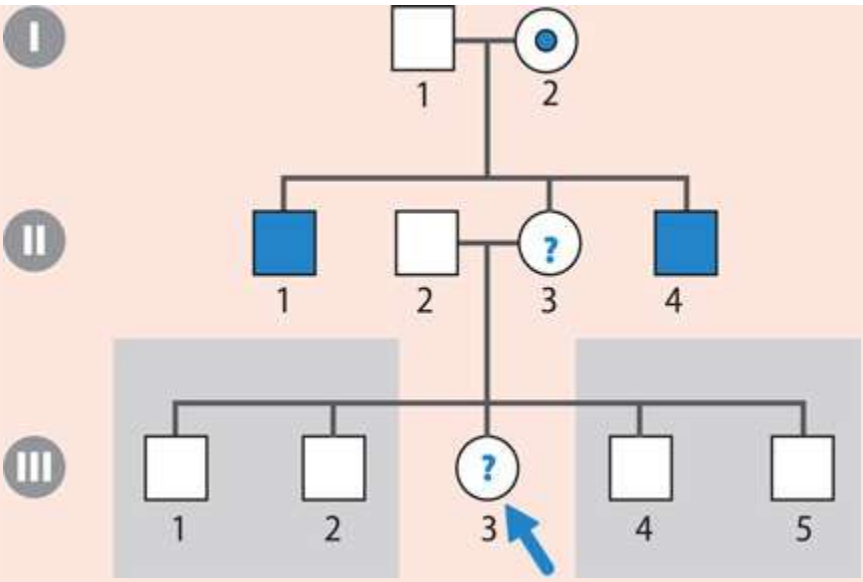


Figure 1 Genetic risk in a pedigree with a childhood-onset X-linked recessive condition. If III-3 were an only child, her risk of being a carrier would be 1 in 4 (her mother, II-3, is the daughter of obligate carrier I-2 and has a 1 in 2 chance of being a carrier; if so, III-3 also has a 1 in 2 risk of inheriting the mutant allele). However, III-3 subsequently mentions that she has four grown-up brothers, none of whom are affected. That additional information suggests that the probability that II-3 is a carrier is much less than 0.5, and that means the chance that III-3 is a carrier is greatly reduced—but by how much?

Bayesian analysis to account for conditional information

If II-3 were a carrier, it would be possible, but unusual, that she would have had four unaffected sons. The new conditional information suggests that she is more likely not to be a carrier, and the risk that III-3 would be a carrier should therefore be much reduced.

The question is: by how much? To answer that question and to give a new risk estimate based on all the information, Bayesian analysis is used. Four steps are involved, as listed below.

1. Identify all the different scenarios that can explain the observations.

2. For each scenario, calculate the prior probability and conditional probability.
3. Multiply the prior probability by the conditional probability to obtain a joint probability for each scenario.
4. Determine what fraction of the total joint probability is represented by each individual scenario to get a posterior probability for each of the three scenarios.

If we discount fresh mutation, there are three possible scenarios in this case: (A) II-3 is not a carrier, and so III-3 is also not a carrier; (B) II-3 is a carrier, but III-3 is not a carrier (because she did not inherit the mutant allele); (C) II-3 is a carrier and III-3 is also a carrier (because she inherited the mutant allele). As detailed in **Figure 2**, Bayesian analysis suggests that scenario A is by far the most likely—the ratio of the probability for the three scenarios is 32:1:1 for A:B:C. Coming back to the original question, the probability that III-3 is a carrier is given by the posterior probability for scenario C, which is 1/34 (or close to 3 %), substantially less than the prior probability of 25 %.

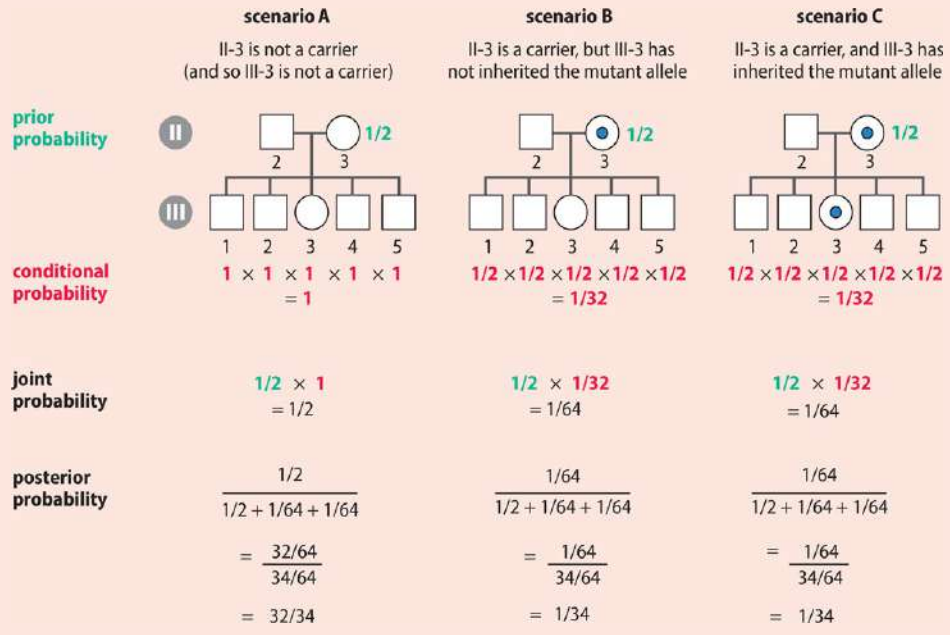


Figure 2 According to Bayesian calculations, the risk that III-3 in Figure 1 is a carrier is only about 3 %. Here, the prior probability is the standard risk due to

Mendelian segregation, and the conditional probability is the *multiplicative product* of the individual probabilities that individuals in generation III have the status that is attributed to them. The probability that an individual male in generation III is unaffected is 1 in 2 if II-3 is a carrier (scenarios B and C), or 1 if II-3 is not a carrier (scenario A). The probability that III-3 is not a carrier is 1 in scenario A (because her mother is not a carrier), or 1 in 2 when her mother is a carrier (scenarios B and C). The joint probability is the product of the prior probability and conditional probability, and the posterior probability is the fraction of the total joint probabilities (for all scenarios) that is attributable to one scenario.

Preimplantation genetic testing is carried out to prevent the transmission of a harmful genetic defect using *in vitro* fertilization

Preimplantation genetic testing is a technique used to identify or screen for genetic defects in embryos created through *in vitro* fertilization before pregnancy so that an apparently healthy embryo can be implanted into the uterus. The procedure is technically challenging because it typically involves analyzing a single cell (as a way of monitoring the genotype of the oocyte or of the early embryo), and is not widely available.

To infer the genotype of an oocyte, polar bodies are sometimes analyzed. More commonly, a single cell (blastomere) is removed from the very early embryo for testing (**Figure 11.18**). For technical reasons, some centers prefer to analyze a few cells taken from the outer trophectoderm at the later blastocyst stage (the trophectoderm will give rise to extra-embryonic membranes). In either case, the remaining embryo can be implanted successfully and is viable.

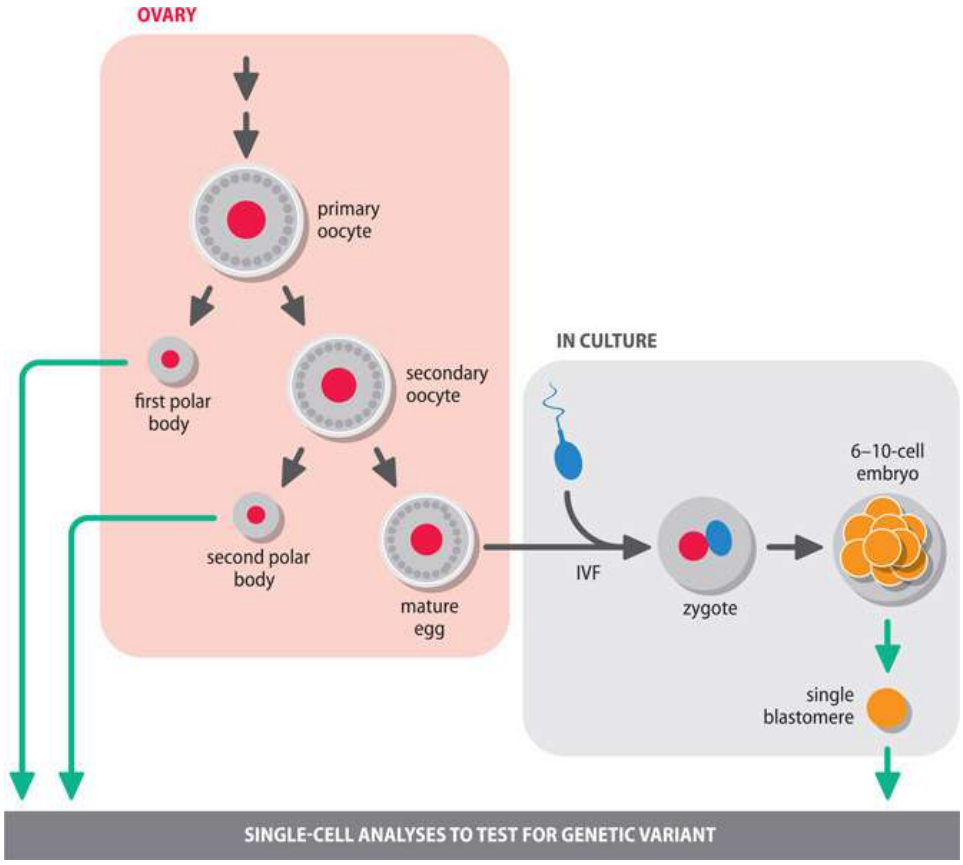


Figure 11.18 Preimplantation genetic diagnosis often involves analyzing single cells. Unlike for sperm cells, the meiotic divisions giving rise to an egg cell are asymmetric: the primary oocyte divides to give a secondary oocyte and a polar body, and the secondary oocyte divides to give the mature egg cell and a second polar body. The polar bodies are disposable and can be analyzed to infer whether the egg cell is carrying a specific harmful genetic variant or a chromosomal aneuploidy. If not, IVF proceeds with what appears to be a normal egg cell. More commonly, a single cell is sampled from the early embryo and tested for the presence of the harmful genetic variant. If the test result is negative, the remaining embryo is implanted in the uterus, and development can proceed normally. Because it can be challenging to obtain data from a single cell, some centers prefer to allow the embryo to develop further and remove a few cells from the blastocyst for testing.

Standard assisted reproduction techniques are used to obtain embryos for testing: ovarian stimulation (to produce eggs that are then collected under sedation), addition of sperm, and assessment of the *in vitro* fertilization

(IVF) and of the embryos produced. In the case of single blastomere analyses, individual embryos are grown in culture to reach the 6–10-cell stage. At this stage a small hole is made in the zona pellucida and a single cell is removed through the hole for testing. Despite the loss of one cell for analysis, the embryo will go on to develop normally.

There are two broad categories of preimplantation genetic testing (PGT), as listed below.

- *Diagnosis*. This applies to couples who are at risk of transmitting a specific genetic abnormality: one or both parents have previously been shown to carry a pathogenic variant or chromosome abnormality that the test is designed to identify.
- *Screening* is performed on couples who may have difficulty conceiving but have no *known* genetic abnormality. Here, the embryo is screened for the presence of any chromosomal aneuploidy.

In both cases, the object is to implant normal embryos only, to avoid the birth of an affected child (in diagnostic cases) or to improve the pregnancy success rate (in screening cases).

For preimplantation genetic diagnosis, prior identification of mutant alleles in one or both parents allows a test in which one or more relevant DNA regions in the DNA from the biopsy are PCR-amplified and sequenced.

If there has been difficulty in identifying a parental mutation, indirect genetic linkage tests can be conducted using a well-established set of polymorphic markers that span the disease gene locus. Occasionally, the test seeks to identify the transmission of a chromosomal abnormality and involves interphase FISH.

The process of achieving a pregnancy becomes medicalized (with potential side effects associated with ovarian hyperstimulation). Additionally, the likelihood of a successful pregnancy outcome is quite low: it is only about 1 in 5 at the start of an IVF treatment cycle (sometimes no

embryos are suitable for transfer, depending on the number of eggs fertilized, and the number and quality of unaffected embryos), but increases to 1 in 3 after embryo transfer.

Noninvasive prenatal testing (NIPT) and whole genome testing of the fetus

Short fragments of cell-free DNA, both fetal and maternal, are present in maternal blood. The fetal DNA fragments arise from placental cells undergoing apoptosis; the maternal DNA fragments originate from the occasional degradation of the mother's cells through apoptosis and necrosis. The fetal DNA fragments are in the minority, accounting for around 10–15 % of the total cell-free DNA in the maternal circulation between 10 and 20 weeks of gestation. Analysis of cell-free DNA in a maternal plasma sample can therefore be sufficient to investigate the genetic composition of the fetus. Obtaining maternal blood is of course somewhat invasive; the non-invasive terminology is used comparatively, given the much higher degree of invasiveness—and risk—that is linked to amniocentesis or chorionic villus sampling.

Because the cell-free DNA in maternal plasma is dominated by maternal DNA, the easiest fetal DNA sequences to identify are those inherited exclusively from the father (they can be readily amplified and detected). That includes Y-chromosome DNA sequences, and noninvasive fetal sexing is now routinely available with a sensitivity of about 90 % and a specificity of 98 %. Testing has also been possible for other exclusively paternal sequences in certain situations (see **Table 11.6** for some applications).

TABLE 11.6 APPLICATIONS OF NONINVASIVE TESTING FORVARIOUS GENETIC CONDITIONS

Genetic condition	Noninvasive testing/diagnosis

Genetic condition	**Noninvasive testing/diagnosis**
Serious X-linked recessive disorders	Fetal sexing test: identifying a female fetus avoids need for subsequent invasive prenatal diagnosis with associated miscarriage risk, but not for a male fetus
Congenital adrenal hyperplasia (21-OH deficiency)	Fetal sexing test: abnormal androgen production in affected female fetuses results in virilization of the external genitalia. After early identification of a female fetus, the fetal adrenals can be suppressed by the oral administration of dexamethasone to the mother. *CYP21A2* haplotype testing can now be offered from 8 weeks via NIPD if pre-pregnancy work-up (from both parents and previously affected child) suggests informative *CYP21A2* haplotype testing
Hemolytic disease of the newborn	Testing for paternal rhesus D blood group: rhesus D-negative women may be at increased risk of hemolytic disease of the newborn (because of a previous affected pregnancy or raised antibody titer). If a paternal rhesus D is identified, the pregnancy needs to be monitored closely because of the risk of fetal anemia
Cystic fibrosis	Haplotype testing can be offered if DNA is available from both parents with confirmed mutation and DNAfrom previously affected child or confirmed non-carrier child. Mutation testing for *CFTR* is also possible if father is known to be carrier of particular *CFTR* variant
Various craniosysnostosis syndromes	Such as Aperts syndrome, Crouzon syndrome, and achondroplasia where a parent is known to have a pathogenic *FGFR2/3* variant
Duchenne/Becker muscular dystrophy	Dystrophin haplotype testing where familial mutation is known

Genetic condition	Noninvasive testing/diagnosis
Spinal muscular atrophy	Where both parents are known to be carriers

Technological breakthroughs

It is technically easy to test cell-free DNA in maternal plasma for the presence of exclusively paternal DNA sequences. More comprehensive testing has been difficult as fetal markers are not readily distinguished from maternal homologs, and there is a large background of circulating maternal DNA in maternal plasma. Technological advances in noninvasive prenatal testing and screening over the last decade have dramatically opened up this field, offering an exciting new window on fetal diagnosis and fetal screening, with the list of current possibilities likely to expand rapidly.

A major breakthrough came from overcoming technical obstacles to what is a very simple principle: counting the parental haplotypes. For any very short genome region, three haplotypes exist in the freely circulating DNA in maternal plasma: the maternal haplotype that is transmitted to the fetus (M_t), the maternal haplotype that is not transmitted to the fetus (M_u), and the paternally transmitted haplotype (P_t).

Because the DNA in maternal plasma will be a mix of DNA originating from degraded maternal cells plus a relatively small amount of fetal DNA, typing for individual DNA markers will show a small excess of alleles from M_t haplotypes over alleles from M_u haplotypes (Table 11.7). To detect such a small difference reliably, a very specific test would be needed; however, with massively parallel DNA sequencing it is comparatively easy to count millions (or even billions) of DNA molecules—permitting very specific testing.

TABLE 11.7 DISTINGUISHING BETWEEN THE TWO MATERNAL HAPLOTYPES BY COUNTING ALLELES AT MATERNALLY HETEROZYGOUS MARKER LOCI IN

DNA FROM MATERNAL PLASMA		
Contribution made by DNA from:	**Expected count of alleles at marker loci on:**	
	Transmitted maternal haplotype, M_t	**Untransmitted maternal haplotype, M_u**
Mother (M_t+M_u)	N(1 –ε)	N(1 –ε)
Fetus (M_t+P_t)	Nε	0
Total	N	N(1 –ε)

ε is the fraction of the DNA in maternal plasma that originates from fetal cells. N is the number of haplo-types with the allele of interest that have been analyzed.

NIPT can be used with targeted massively parallel DNA sequencing: certain genome regions of interest are captured from the genomic DNA (see Box 11.2 above for the principle) and sequenced.

One application of NIPT is fetal aneuploidy screening. Previously, the problem here had been distinguishing fetal autosomes from the maternal equivalents. When fetal DNA accounts for 10 % of the DNA in maternal plasma, the amount of chromosome 21 DNA increases by just 5 % if the fetus has trisomy 21. Because of massively parallel DNA sequencing, that small difference can be readily detected (trisomy 21 can be detected with a sensitivity of 99 % and a specificity of 99 % using this method). Current evidence suggests that NIPT is more cost-effective as a screening tool (to define a high-risk group that can then be offered confirmatory amniocentesis) than as a diagnostic procedure.

An overview of the different types of genetic screening

The genetic testing described above is reactive: it is carried out in response to individuals seeking medical help or advice about the risk of developing or transmitting a genetic disorder. That is to say, it comes to the attention of medical services as a disease phenotype in a family member. A causative genotype is then sought so that other family members at risk can be tested to see if they have the pathogenic variant.

In *genetic screening*, the genetic tests are carried out in communities and populations and genotypes are used to predict phenotypes that may manifest at some future time. In population screening, a particular population is screened with a targeted enquiry.

With the advent of whole genome sequencing, there is an added dimension to the screen; the whole genome assay is in effect a screen, to which targeted enquiries are then made depending on the clinical question.

Genetic screening of populations can be carried out using biochemical and physiological markers (as products of genetic variants) as well as looking directly at genetic make-up. In Section 8.3 we considered a type of longitudinal population screening, exemplified by the UK Biobank project, in which comprehensive testing is carried out on people at regular intervals over decades. That is a research-led type of screening without any (or only very exceptional) feedback of findings to participants. In contrast, the three types of genetic screening listed below are primarily directed at providing clinical benefit to the subjects tested.

- *Pregnancy screening*. The object is to identify whether or not the pregnancy is at a very high risk of leading to the birth of a child with a serious genetic disorder. The motivation for the test is usually to prevent the birth of an affected child. (Less commonly, the test might be requested to allow psychological preparation and medical management planning for such a birth, while also offering psychological benefit, should the test indicate that the fetus is unaffected.) It may entail screening for a serious single-gene disorder in communities where that disorder is prevalent, or for aneuploidies. As described below, technological advances now permit comprehensive genetic profiles to be obtained for a fetus. That might lead to new ways of treating disease *in utero*.
- *Newborn screening*. This is carried out in many countries, but to variable extents. A major motivation has been to target early treatment in serious disorders for which early intervention can make a substantial difference and may lead to disease prevention. Genetic

screening of newborns began with certain metabolic disorders, and this class of disorder is still a major focus.

- *Carrier screening*. This is also carried out in many countries, often targeted at particular ancestral groups (for example, Tay-Sachs or sickle cell screening) with an aim of identifying carrier couples of a mutant allele for a range of severe autosomal recessive disorders. More recently, advances in sequencing techniques have led to expanded carrier screening approaches where many different carrier states are screened for simultaneously.

Pregnancy screening for fetal abnormalities

Specific maternal screening programs have been undertaken in the first trimester to identify fetuses at high risk of common and serious single gene disorders—such as sickle-cell disease and thalassemia—that are prevalent in certain communities. However, the focus for most prenatal screening is maternal screening for fetal aneuploidy, notably the commonest chromosomal abnormality, trisomy 21 (causing Down syndrome).

As described above, massively parallel DNA sequencing is increasingly used to screen DNA in maternal plasma (which includes small amounts of DNA from fetal cells) for evidence of aneuploidies such as trisomy 21. This type of NIPT is still largely offered after a "high risk combined" screen result based on three parameters. One is *nuchal translucency*, the skin thickness at the back of the neck, as measured by ultrasound scanning between 11 and 14 weeks of gestation; it is determined by the amount of fluid that collects here (which is often greater in Down syndrome babies). A second factor is the mother's age (the risk increases 16-fold as the maternal age increases from 35 to 45 years). The third factor is based on altered levels of certain maternal serum proteins, such as an increased level of free b-HCG (human chorionic gonadotropin) and a decrease in PAPP-A (pregnancy-associated plasma protein A).

Private clinics are increasingly offering NIPT to a general pregnant population but here the pre-test probability of an aneuploidy is usually less

than 1 %. Whilst this may still be a risk some couples would not wish to tolerate, a quick look at Bayes theorem (Box 11.5 on page 458) tells us that the lower the population frequency, the higher is the chance that a positive result will be a false positive. For example, if a 30-year-old woman has a chance of an aneuploidy of 0.1 % and the NIPT has a sensitivity and specificity of around 99 % (commonly advertised as such), then the chance that her positive test is a true positive is <10 %. For a 1 % figure the probability of a false positive is 50 %. This therefore has the potential to lead to a much higher rate of inappropriate invasive follow-up testing than people commonly realize.

On the basis of combined screening, approximately 2 % of women will have a greater than 1 in 150 risk (compared with an overall population risk of about 1 in 670); they will be offered chorion biopsy for definitive aneuploidy testing. There will be an adverse outcome in 20 % of these women (which includes trisomy 13 and trisomy 18 in addition to trisomy 21); that still means there is no chromosome abnormality in 80 % of women who take up chorion biopsy—the development of a reliable test based on cell-free fetal DNA in the maternal serum is therefore a major advance. The combined screening detects 90 % of all affected pregnancies.

First-trimester ultrasound is important in estimating the date of delivery and for nuchal measurement, and is essential for accurate estimates of gestational age needed for risk calculations based on the levels of the maternal serum proteins described above. Additional ultrasound is routinely offered in pregnancy at around 20 weeks of gestation to look for structural anomalies (a significant proportion of which are due to chromosomal or Mendelian disorders).

Newborn screening allows the possibility of early medical intervention

Newborn screening was pioneered in the late 1960s. Screening for phenylketonuria (PMID 20301677) used dried blood spots collected on a filter-paper card (the Guthrie card) at 5 days of age. Assays for congenital

hypothyroidism (which has a number of causes, few of which are genetic), were added shortly afterward. For both conditions the rationale was prevention of the developmental delay that would inevitably ensue in the absence of medical intervention (which involves dietary changes for phenylketonuria and hormone replacement for congenital hypothyroidism —see **Table 11.8**).

TABLE 11.8 NEWBORN SCREENING PROGRAMS FOR SELECTED AUTOSOMAL RECESSIVE DISORDERS AND CONGENITAL HYPOTHYROIDISM

Genetic disorder	Prevalence	Type of screening	Treatment of affected individuals
Congenital hypothyroidism	1 in 5000	assay of free thyroxine or thyroid-stimulating hormone in serum	hormone replacement
Cystic fibrosis	1 in 2500 in European populations	screen for immunoreactive trypsinogen, then confirm by scan for *CFTR* mutations	antibiotics, chest physiotherapy, pancreatic enzyme replacement for those with pancreatic insufficiency

Data from guidelines proposed by the American College of Medical Genetics and Genomics, which makes information on screening programs for individual disorders available through PubMed (PMID 21938795) and the NCBI bookshelf (http://www.ncbi.nlm.nih.gov/books/NBK55827/).
CFTR, cystic fibrosis trans-membrane conductance regulator gene; HbF, fetal hemoglobin; HPLC, high-performance liquid chromatography; IEF, isoelectric focusing.

Genetic disorder	Prevalence	Type of screening	Treatment of affected individuals
Galactosemia	1 in 75000	assay of levels of erythrocyte galactose-1-phosphate and galactose-1 - phosphate uridyltransferases	change of diet to reduce intake of galactose
Phenylketonuria	~1 in 12000	plasma amino acid analysis to show increased phenylalanine: tyrosine ratio	change of diet to reduce intake of phenylalanine (Clinical Box 9 on page 234)
Sickle-cell disease	~1 in 500 with African ancestry	hemoglobin separation by electrophoresis, IEF, or HPLC. DNA studies may be used to confirm genotype	hydroxyurea (increases HbF in red blood cells, reducing transfusion requirement and decreasing frequency and severity of vaso-occlusive events; prophylactic penicillin

Data from guidelines proposed by the American College of Medical Genetics and Genomics, which makes information on screening programs for individual disorders available through PubMed (PMID 21938795) and the NCBI bookshelf (http://www.ncbi.nlm.nih.gov/books/NBK55827/).
CFTR, cystic fibrosis trans-membrane conductance regulator gene; HbF, fetal hemoglobin; HPLC, high-performance liquid chromatography; IEF, isoelectric focusing.

Inborn errors of metabolism have been a major focus of newborn screening for two reasons. First, they have been studied for decades, and there is a highly developed understanding of the molecular basis of disease, allowing useful early medical interventions in some cases.

The second advantage is that inborn errors of metabolism are typically amenable to easy-to-use screening systems that work at the gene-product or metabolite level, and are applicable to easy-to-access patient samples, such as blood or urine. A disease allele may have any one of a potentially very large number of different mutations; if the gene has many exons, the screening can be laborious. However, all that heterogeneity at the DNA level often has a rather uniform effect at the gene-product level: a single assay can often detect abnormalities in the product or characteristic changes in certain metabolites.

As a result, it is usual to use assays at the gene-product level, or assays for disease-associated metabolites (tandem mass spectrometry—which allows the parallel testing of multiple metabolites in blood and urine samples—can efficiently screen for a range of metabolic disorders at low cost).

Benefits versus disadvantages of newborn screening

More recently, other disorders have been added to screening lists, and the huge advances in massively parallel DNA sequencing have led to proposals to greatly increase the number of disorders that are screened for.

In addition to the large costs of implementing national screening programs, any screening program will include false positives. Anxiety can be generated in families who receive a positive screen result but whose child is unaffected on second testing (and as mentioned above, the more tests that are taken, the greater is the chance of receiving a false positive result). Accordingly, some countries have taken a quite conservative approach. In the UK, for example, national newborn screening is restricted to nine rare but serious conditions: phenylketonuria, congenital hypothyroidism, medium-chain acyl-CoA dehydrogenase deficiency (MCAD), sickle-cell disease, cystic fibrosis, maple syrup urine disease, isovaleric acidaemia, homocystinuria, and glutaric aciduria type 1.

By contrast, the American College of Medical Genetics and Genomics (ACMG) has recommended screening for 54 conditions (including hemoglobin abnormalities, various inborn errors of amino acid, fatty acid, or organic acid metabolism, biotinidase deficiency, congenital adrenal hyperplasia, galactosemia, and cystic fibrosis).

Early treatment might not be of clinical benefit in all of the conditions screened, but there can be other benefits. One benefit might be a greater awareness of the disorders and a greater sharing of information, increasing knowledge of the natural history of these very rare disorders. Another is that parents will be informed about the condition and recurrence risks before they have further children. Some countries have piloted newborn screening for Duchenne muscular dystrophy, not because of therapeutic benefit but because if a child does not present until 4 years of age, couples may already have a second affected child at the time of diagnosis.

Newborn screening using whole genome sequencing (WGS)

The UK government has recently announced plans for newborn screening using whole genome sequencing (https://www.gov.uk/government/publications/genome-uk-the-future-of-healthcare). At the time of writing, Genomics England are leading a public consultation on the possibility of introducing whole genome sequencing as the primary technology through which to offer newborn screening, which offers the potential of diagnosing many more conditions, as well as indicating sensitivities to future pharmacological interventions, in a newborn population. Whilst that may sound appealing, the disadvantages will need careful attention before existing screening programs expand in this way.

As alluded to above, WGS here is a new type of screen in itself; it is the assay upon which the newborn screen is performed, and this can be predetermined by determining which filters are applied to the data. This is similar to the filtering of tandem mass spectrometry outputs used in current

screening programs. That is to say, only certain conditions will be looked for in the data that have the potential to reveal more. Questions arise about who then might access the remaining data for predictions, and when they might do so. For example, will parents have a right to obtain or access the data? Can they—or the data analysts they instruct—inspect it for adult-onset conditions, uncertain findings and so on? These questions about such future predictions often seem qualitatively different when talking about existing data from WGS than data not brought into existence through traditional screening techniques. That said, current newborn screening primarily relies on interpreting the peaks from tandem mass spectroscopy; many more abnormal gene products could be analyzed this way than is currently routine.

Patient support organizations such as the Genetic Alliance UK are cautiously supportive of whole genome approaches to newborn screening, but have called for careful scrutiny of the ethical issues involved.

Different types of carrier screening can be carried out for autosomal recessive conditions

Carrier screening can be carried out at the pre-conception level or antenatally, but it can also sometimes be a "side-effect" of newborn screening (in that newborn screening does not aim to detect carrier status that is not relevant to an individual until of reproductive age). Ideally carrier screening is done when couples are planning to have children, to see whether they are both carriers of the same condition, but often the window of opportunity for this may be small or opaque.

The example of β-thalassemia screening

Approximately 70 000 babies are born each year with this disorder, the incidence being highest in Mediterranean countries, India, Africa, Central America, the Middle East, and Southeast Asia. To treat the resulting anemia, affected individuals require repeated blood transfusions; however,

that causes iron overload, which in turn leads to liver damage and cardiomyopathy. Iron chelation therapy is then used to increase iron excretion, prolonging life expectancy well into the fourth decade of life and usually beyond that.

Carrier screening can be undertaken using mean corpuscular volume and mean corpuscular hemoglobin levels in the standard full blood examination; various methods are used to confirm the diagnosis. In 1973, carrier testing was introduced across Greece and Cyprus after educational programs at schools, in the armed forces, maternity clinics, through the mass media, and, in Cyprus, through the Orthodox Church. Sardinia introduced screening a few years later.

Subsequently, many countries have developed screening programs. In Iran, many provinces of Turkey, the Gaza Strip, and Saudi Arabia the testing is mandatory for couples registering for marriage. In the Gaza Strip, couples have to sign a declaration that they are aware both are carriers if they continue with the marriage. These countries have opted for screening before pregnancy; in other countries, screening occurs in antenatal clinics—if the woman is found to be a carrier, testing is offered to her partner. Although consent is an intended prerequisite, screening evaluations have indicated that patient awareness and understanding of the program are very variable.

There has been a significant reduction in affected births in countries with screening programs, partly due to altered marriage plans but mainly due to the uptake of prenatal diagnosis and termination of pregnancy. For example, the incidence of b-thalassemia in Sardinia when screening was introduced in 1975 was 1 in 250 births; by 1995 it was 1 in 4000 births. In Cyprus the number of affected births in 1974 was 51, in 1979 it was 8, and there were no affected births between 2002 and 2007. Similar marked reductions have been reported after the introduction of antenatal screening programs in Taiwan and Guangdong China.

The example of Tay-Sachs disease screening

Carrier screening programs have also sometimes been directed to particular population groups with a high incidence of a serious disorder. For example, the recessive disorder Tay-Sachs disease (PMID 20301397) is a progressive neurode-generative disorder that is rare in most populations (with a carrier frequency of about 1 in 300 in Europe and America), but is especially common in Ashkenazi Jews (about 1 in 27 is a carrier). This inborn error of metabolism presents with progressive weakness and loss of motor skills at between 3 and 6 months, followed by seizures, blindness, spasticity, and death usually before 5 years of age. It is caused by failure to produce the enzyme hexosaminidase A, as a result of genetic mutation in the *HEXA* gene. As a consequence, a fatty substance, GM2 ganglioside, accumulates in brain cells and nerves, damaging and eventually destroying them.

Carrier testing based on assaying serum hexosaminidase A began in 1970, when it was recognized that carriers may be distinguished from non-carriers by this assay. Testing is available through health services in many countries; the Dor Yeshorim organization also offers genetic screening to Ashkenazi Jews worldwide through orthodox Jewish High Schools and in community testing sessions. When testing is undertaken in orthodox schools, the results may not be given directly but instead be available at a later stage for couples considering marriage. This screening program has led to a significant reduction in the number of children born with Tay-Sachs disease in this community.

Preconception couple carrier screening

The object of preconception couple carrier screening is to identify couples who are each carriers of a pathogenic variant for the same severe autosomal recessive disorder (which can be any from a range of such disorders). Such screening programs have been in place over the past several decades in communities where a particular disease has a high prevalence (for example, screening for Tay-Sachs or sickle cell disease). With the advent of faster, cheaper, whole genome approaches (and the realization that everyone is a

carrier of roughly 1–10 autosomal recessive conditions), "*expanded pre-conception screening*" (ECS) approaches can offer simultaneous screening for say, 100, serious recessive conditions, regardless of population prevalence.

The chance of any one recessive condition may be extremely low, but the combined chance of finding a couple where both members are carriers for the same recessive condition is around 1 % in unselected populations. Innovative approaches that have been piloted in the general population of the Netherlands, and elsewhere, have disclosed only couple results; that is to say, prospective parents are told in cases where they both test positive for the same disease carrier state, and individual carrier results are not disclosed on the basis they will have no medical consequences. Such expanded screening can also be done in early pregnancy but at this point reproductive options are more limited (termination or not) than if preconception screening is carried out.

New genomic technologies are being exploited in cancer diagnostics

As sequencing technologies have improved in depth as well as breadth, they have played a crucial role in elucidating cancer mechanisms. Single-cell sequencing is helping to define the evolution of cancers, and the complex relationships between different cancer subclones is being defined over space and time, demonstrating the enormous heterogeneity of cancers and the difficulty of successfully treating them. Genetic and genomic technologies have also driven improvements in testing for cancer in different ways.

Diverse cancer biomarkers

Different genes associated with certain types of cancer can provide **biomarkers** of those cancers that when detected provide clinically useful information about the cancer. Different types of nucleic acid biomarker can

be found, including alleles with a specific pathogenic point mutation, oncogenic fusion genes, and specific gene expression signatures. Information from a detected biomarker can be used in different ways: to diagnose a cancer phenotype; to predict the likely response of the cancer to defined drugs; to indicate the likely clinical course of the cancer; and finally, to monitor the cancer (assessing the presence of mutant clones, and so on). **Table 11.9** gives examples of some of the very many different biomarkers used in cancer testing.

TABLE 11.9 EXAMPLES OF DIFFERENT ROLES FOR DNA AND GENE EXPRESSION BIOMARKERS IN CANCERTESTING

Role	Gene/expression biomarker	Cancer type (comment)
Diagnostic	*BCR-ABL1*	chronic myeloid leukemia (see Figure 10.8A)
	JAK2	myeloproliferative disease (specific mutations confirm diagnosis of clonal disease)
	EWS-FLI1	Ewing sarcoma
Predictive	*HER2*	breast cancer (amplification predicts response to anti-HER2 antibodies)
	BRAF	melanoma (specific point mutations predict response to specific BRAF inhibitors)
	KIT,PDGFRA	gastrointestinal stromal tumors (specific point mutations predict response to c-KIT/PDGFRA inhibitors)
Prognostic	*TP53*	chronic lymphocytic leukemia (specific point mutations are indicative of poor outcome)

MRD, minimal residual disease.

* Multi-gene expression signatures.

Role	Gene/expression biomarker	Cancer type (comment)
	BRAF	metastatic colorectal cancer (specific point mutations are indicative of poor outcome)
	* *MammaPrint* (70-gene)	breast cancer (risk stratification)
	* *OncotypeDx* (21 -gene)	breast cancer (risk stratification)
Disease monitoring	*BCR-ABL1*	chronic myeloid leukemia (detection of MRD)
	PML-RARA	acute promyelocytic leukemia (detection of MRD)

MRD, minimal residual disease.

* Multi-gene expression signatures.

Multiplex testing using targeted DNA sequencing

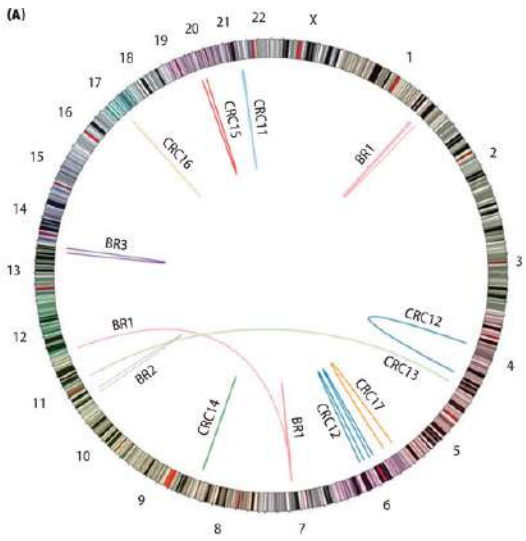
As described in Box 11.2 on page 441, targeted DNA sequencing allows DNA sequences from any genome region of interest to be selectively captured and sequenced. Multiplex testing for panels of cancer susceptibility genes have now been adopted by many diagnostic services. There has been close liaison between oncology and genetic services to determine which cancer diagnoses have a significant germline predisposition, how this might affect treatments, and subsequent cascade testing of family members.

In the UK, the NHS genome medicine service has more recently delineated the aim to provide a uniform cancer testing service for all cancers, and current panels for cancer (and for rare and inherited disease) are listed in the NHS National Genomic Test Directory at: https://www.england.nhs.uk/publication/national-genomic-test-directories/

Noninvasive cancer testing uses “liquid biopsies”

Another promising recent development is noninvasive cancer testing. Instead of taking a tumor biopsy (which can be difficult, according to the type of cancer), different approaches allow the analysis of freely circulating tumor DNA in plasma, as described in Clinical Box 14 on page 412. (They were stimulated by the application of high-throughput DNA sequencing in noninvasive prenatal testing of the fetal genome, as described above.)

The freely circulating DNA originates from cells undergoing apoptosis or necrosis, which includes originally healthy cells as well as inflamed cells and diseased cells, such as cancer cells. This means that tumor-specific variants need to be detected against a background of circulating DNA from non-tumor cells in the same individual.



(B)

Sample	Chromosome	Type of change
CRC11	21	inversion
		deletion
		deletion
CRC12	4	inversion
	6	deletion
		deletion
CRC13	11 and 4	interchromosomal
CRC14	8	duplication
CRC15	20	inversion
CRC16	17	inversion
CRC17	6	inversion
BR1	1	deletion
		inversion
	7 and 12	interchromosomal
		deletion
BR2	11	deletion
BR3	13	inversion

Figure 11.19 Detection of tumor-specific rearrangements by massively parallel DNA sequencing in plasma samples. (A) A Circos plot indicating rearrangements identified in tumor-cell DNA present within plasma samples from 10 cancer patients, 7 with colorectal cancer (CRC11 to CRC17), and 3 with breast cancer (BR1 to BR3). No rearrangements were identified in DNA obtained from plasma samples from 10 unaffected controls. (B) Observed DNA rearrangements. A table with details of the breakpoint coordinates is found in the original paper. Note that droplet-digital PCR methods, described in Section 11.2, can permit very sensitive quantitation of copy

number variation (Adapted from Leary RJ et al. [2012] *Sci Transl Med* 4:162ra154; PMID 23197571. With permission from the AAAS.)

Massively parallel DNA sequencing can be applied to the analysis of plasma DNA. As well as detecting sequence variants it can identify tumor-specific chromosome alterations with comparative ease (**Figure 11.19**). Whilst this represents another technological breakthrough, the use in clinical practice will depend on the ability to detect cancers accurately at a more treatable stage. Different methods can be used to quantitate mutant alleles and copy number variants, but droplet digital PCR (described near the end of Section 11.2) can be particularly useful. Because of its high sensitivity, it can rapidly detect *minimal residual disease* (the small number of cancer cells that remain in the body after treatment that may become active, start to multiply and cause a relapse).

Bypassing healthcare services: the rise of direct-to-consumer (DTC) genetic testing

Up to this point we have considered genetic testing offered through healthcare services. That was the only option until recently, when commercial genetic testing began to be offered directly to consumers. Two stimuli in particular have led to the growth of DTC genetic services: the recent rapid decrease in the cost of genetic testing, and increasing identification of genetic variants conferring susceptibility to diseases that are common in populations.

The purpose of DTC genetic testing is quite different to that of healthcare-led genetic testing. Genetic testing organized through healthcare services is targeted to people at high risk of developing specific genetic conditions and serves to explain and/or manage health problems. By contrast, DTC testing targets healthy people and the rationale is to facilitate life planning. The development of DTC testing is taking place in the context of a public discourse about personalized/precision medicine and genetics that tend to enthusiastically promote it in a very optimistic light, rarely dwelling on potential concerns and limitations. Such an over-optimistic

perspective potentially raises inappropriate expectations of our ability to interpret what common genetic variants mean for our health.

Most DTC genetic tests rely on inexpensive SNP-chip genotyping, which checks for the presence or absence of single nucleotide polymorphisms (SNPs), or small insertions or deletions throughout a genome. SNP-chip genotyping detects common genetic variants well, but detection of very rare variants is poor: "calls" of rare cancer predisposing variants, such as pathogenic variants in *BRCA1* and *BRCA2* and those causing Lynch syndrome, are often false positives.

Genome sequencing is another method becoming more widely used in DTC genetic tests. These tests are not so vulnerable to miscalling as SNP arrays are. However, detecting variants is not the same as knowing their clinical effects—as described above, interpretation of genetic variants can be challenging and often difficult in the absence of a (familial) phenotype.

DTC tests currently sit outside much of the regulation governing clinical genetic testing, but claim to provide insight into issues as diverse as ancestry, nutrition, athletic ability, and child talent. Many testing providers also claim to help provide insight on health. The range of health information that might *potentially* be provided could include the items listed below.

- *Polygenic risk scores*—Combining many different common variants across the genome may serve to place someone in a broad risk category, such as: "your genes predispose you to weigh about 3 % more than average". The validity and utility of these risk scores for predictive clinical purposes is hotly debated. In our opinion, although polygenic scores may be useful in researching the causes of disease, or stratifying populations into higher and lower risks, they are rarely able to usefully predict disease—see the end of Section of 8.2 for more detail.
- *Genotyping at specific points*—looks at specific variants that influence the chance of developing particular diseases, such as: "you have two copies of the e4 variant in the *APOE* gene. People with

this result have an increased risk of developing late onset Alzheimer disease." This type of testing can also be used to identify variants that affect drug metabolism.

- *Carrier screening*—looks at specific variants to identify people who are carriers for particular recessive genetic conditions, such as: "one variant detected in the *CFTR* gene. If you and your partner are both carriers, each child may have a 25 % chance of having this condition." Many carrier tests are ancestry specific: they test for specific carrier variants common in a particular ancestral group. If someone with a different ancestry were a carrier, this would probably not be detected because it would likely be due to a different variant (which the test would not check).
- *Uninterpreted "raw" genetic data*—See below, under DIY genetics, some DTC genetic test companies provide access to uninterpreted genetic data. Customers can download their data and seek an interpretation using third party services. These usually work by cross-referencing the data against freely available genetic databases and by constructing a report based on interpretations in these databases (which may not be up to date). They may report variants and disease risks that were not reported, or referred to, by the original DTC genetic test company, and might repurpose raw data from tests designed to answer other questions, such as ancestry, to try to provide health information.

In reality, the health information provided by many DTC companies is far from comprehensive. For example, a recent analysis of 15 DTC genetic testing companies advertising to UK consumers found that none of them complied with all the UK Human Genetics Commission's principles for good practice regarding consumer information. The "personalized medicine" that genetic testing promises is often portrayed in an optimistic light by the mainstream media, and genetic technology is generally presented as highly accurate. As a result, people may perceive genetic

testing as clearly predictive, and expect that the results will help them plan for the future.

DIY genetics

Do-it-yourself (DIY) genetics has also risen in popularity. Here, people ask for raw data from DTC companies and process this themselves via third-party interpretation services. But many variants that are called in the raw data of a DTC test and sent for clinical confirmation are false positives. This limitation is often not appreciated by DTC customers, or the health professionals they may subsequently visit, leading to anxiety and inappropriate medical interventions.

The limitations of DTC genetic tests

The predictive value of these tests is often low when there is no family history of disease. A person with no medical or family history of disease X is informed that "you have a disease X-causing (or 'disease X-predisposing') genetic variant." It may be that there are currently unmeasurable protective genetic (or other) factors in that person's family that mean that the variant is less likely to lead to disease X in that person.

Even if a person does have a family history, identifying a "high genetic risk" via DTC genetic testing does not mean that they will definitely develop the condition. A study of people with a genetic form of diabetes found that up to 75 % of those who carry a particular missense variant in the *HNF4A* gene—specifying an R114W substitution—developed diabetes by age 40. However, a recent study looking at the same variant in UK Biobank participants who were not pre-selected as having diabetes showed that only 10 % developed diabetes by age 40.

False positives are common, especially when SNP-chip genotyping and third-party interpretation services are used. "Miscalls" due to inaccurate

genotyping of rare variants when using SNP-chips can be common: a recent study using SNP-chips to genotype very rare pathogenic *BRCA1* and *BRCA2* variants in UK Biobank participants found that 96 % were false positives. Additionally, the databases used by third party interpretation services to interpret the data may not be up to date, and variants may be classified incorrectly due to outdated evidence.

False negatives are another significant concern. This can happen because DTC genetic tests tend to prioritize breadth over detail. At the time of writing, for example, the 23andMe "genetic health risk" report for *BRCA1* and *BRCA2* currently only checks for just three disease-causing variants, which are mainly relevant for people with Ashkenazi Jewish ancestry. But there are thousands of different pathogenic *BRCA* variants that the test does not check for. As a result, about 80 % of people with disease-causing *BRCA* variants in the general population might be expected to be given false reassurance that their *BRCA1* and *BRCA2* testing was negative.

The downsides of improved sensitivity through whole genome sequencing: increased uncertainty about what variants mean

The prior probability of any one variant identified via whole genome sequencing (WGS) being causative for a patient's rare disease is extremely low. Attempts to catalogue human genetic variation, for example via the 1000 Genomes Project, show that a typical human genome differs from the reference human genome at between 4 and 5 million sites. Most of these variations will be benign, some may subtly impact on risk of various common diseases, and a very small number will have the potential to cause serious disease.

Careful filtering of WGS datasets is therefore crucial to produce a meaningful output. This in turn requires a significant change in mindset from the one that prevailed in the days when most pathogenic variants were identified after phenotype-driven single-gene sequencing. In those cases, identified variants would have a much higher prior probability of being causative. Now, in the complex world of WGS datasets, variants should be

"innocent until proven guilty". Translating this principle into clinical practice, however, is difficult in the context of a prevailing view that sees a genome sequence as a "blueprint" that speaks for itself. In practice, extensive filters are often applied by bioinformatic pipelines, to create virtual gene panels (as described in Section 11.3), which means that only certain variants stand a chance of becoming a communicated result.

Even when variants are consistently picked out through a pipeline, there remains considerable discrepancy in how different laboratories interpret the same variant. One study showed how laboratories using the same guidelines agreed on their classification in just one-third of cases; about a quarter were classified so differently that different medical interventions would be recommended. International guidelines for variant interpretation are helpful but it is arguably very unlikely that they will ever reach a stage where the outputs from WGS can be easily filtered to produce clinically meaningful results for a person without the need for additional expertise that links (familial) phenotype with genotype.

Improving knowledge of variant interpretation also leaves us with a difficult legacy, that some patients will have been diagnosed incorrectly with genetic conditions, yet healthcare services to date have no systematic attempts to revise the data and recontact patients seen in the past. Furthermore, there is often no threshold for communicating genetic variation of uncertain significance (VUS). There is some evidence that people misinterpret these as being definitely pathogenic or definitely benign. However, there is also evidence that many people are uncomfortable with the idea that decisions about non-disclosure might be made without involving them.

Genomic data repositories linked to phenotypes are expanding, but, as of 2022, there is a significant skewing of data from people of European ancestry. Lack of diversity in populations that contribute to genomic data has been linked to both missed, as well as false positive, diagnoses of disease in populations of non-European ancestry. Although there have been many calls to improve the diversity of genomic data sets, and indeed reference genomes utilised, distributing the benefits of genomic research to

all populations needs to address the needs of hitherto underserved communities who may already feel disenfranchised by developments in genomics that are of little apparent utility to them.

11.5 ETHICAL, LEGAL, AND SOCIETAL ISSUES (ELSI) IN GENETIC TESTING

We have described the medical advantages of genetic testing and changing landscape through rapidly evolving genetic technologies. In this section we consider the ethical, legal, and societal issues raised by these developments, often abbreviated as ELSI.

For monogenic disorders, or disorders where there is a single genetic factor that explains most of the risk, close relatives and potential offspring of persons who have inherited (and therefore "carry") the relevant mutant alleles (or genetic variants) may be at high risk of developing the disorder. Positive identification of harmful genetic variants in one person therefore raises the stakes for unaffected relatives who may subsequently be found to have inherited the same genetic variants. As we enter an age where genetic testing—even for single gene disorders—creates a data resource on an individual's entire genome, considerations of how such data should be stored and accessed to utilize their potential to predict current or future health issues (for both the person tested, and their biological relatives) are urgently needed.

Genetic information as family information

The familial nature of genetic information often generates discussion on confidentiality issues. The confidentiality of individual patients should be respected. But we also need to ensure that their close relatives have access to information on possible inheritances that may be relevant for their own health and life choices. Clinical guidance in this area has increasingly taken the stance that genetic information should be seen as confidential to families, not individuals—although for a given individual the personal

consequences of having a genetic change, that is, the clinical symptoms that arise from it, should be confidential to them alone. While patient confidentiality is important in genomics, as in other areas of medicine, the duty to maintain this confidentiality is not absolute: it must be balanced against others, such as the duty to prevent serious harm.

Utilizing genetic information obtained for one person to provide accurate testing for a relative is different from informing a family of the particular details of an individual patient's medical problems. Research indicates that patients often see genetic information as belonging to their family rather than exclusively to them. Healthcare professionals, however, are often reticent about taking a family-centered approach to the confidentiality of genetic information in practice. They worry that this stance could disrupt family dynamics or erode patient trust in the health service (see Dheensa et al. [2017] under Further Reading).

Health professionals need to recognize the nuance above. In many cases, a default approach of not disclosing any information without written consent from specified people, is not "playing it safe" from a legal perspective; it is contrary to professional guidance and vulnerable to legal challenge. When faced with uncertainty about whether to disclose information, health professionals should undertake (and document) a balancing act, considering whether in this instance the duty to preserve individual patient confidentiality is, or is not, outweighed by the potential benefits of disclosure to a patient's wider family. **Box 11.6** reports a recent UK court case centered around the personal versus familial nature of genetic information.

CLINICAL BOX 18 CASE STUDY: CONSENT VERSUS DUTY OF CARE TO FAMILY MEMBERS

The ABC vs St George's Case (https://www.bailii.org/ew/cases/EWHC/QB/2020/455.html) involved a patient at St George's Hospital who, on the grounds of diminished responsibility, had been convicted of killing his wife, and been detained in

a secure unit. While there, he developed signs and symptoms consistent with Huntington disease (HD), but he did not want any of his family members to know this. The doctors noted that the man's daughters each had a 50 % chance of inheriting the condition from him, and when informed that one daughter was in the early stages of an unplanned pregnancy, they discussed whether to tell her of her risks. Although they thought she had a right to know, they felt they could not tell her this without her father's consent, which he steadfastly withheld.

After the daughter had given birth, one of her father's doctors told her, by mistake, about the HD diagnosis. She subsequently had a genetic test and found she had inherited the condition. She claimed her father's doctors were negligent in withholding this information from her. She argued that, had she known she would develop the condition herself, and be unable to raise her child, she would have had a termination of pregnancy (regardless of whether the baby had inherited the condition or not).

Although the case for negligence was dismissed, the ruling made it clear that doctors in this situation had a duty to weigh the daughter's interests in the balance, and that using the absence of her father's consent as a veto was misplaced. The ruling confirmed the professional guidance that was already in place (which recommended a breach of confidence where the harms to identifiable others were significant, and which used familial genetic examples to illustrate this). Professional guidance also urges doctors to distinguish a breach of medical confidentiality (the fact that he had a diagnosis of HD) from disclosure about the familial risk (the fact that signs, symptoms, and family history could suggest a genetic condition). Whilst the daughter might then infer she inherited HD from her father, this inference would not equate to a breach of his confidence.

Genetic testing in one person may thus raise a series of ethical (and legal) questions about current and future family members (see **Figure 11.20**). They include:

- Whose responsibility is it to alert relatives of their risks, and how might this best be facilitated?
- How can these familial aspects be appropriately covered in any consent process?
- What does it mean to respect confidentiality when a result might indicate others are at risk?
- When is the risk of a future condition a legitimate reason to terminate a pregnancy?
- Is it appropriate to test children for genetic conditions unlikely to manifest until adulthood?

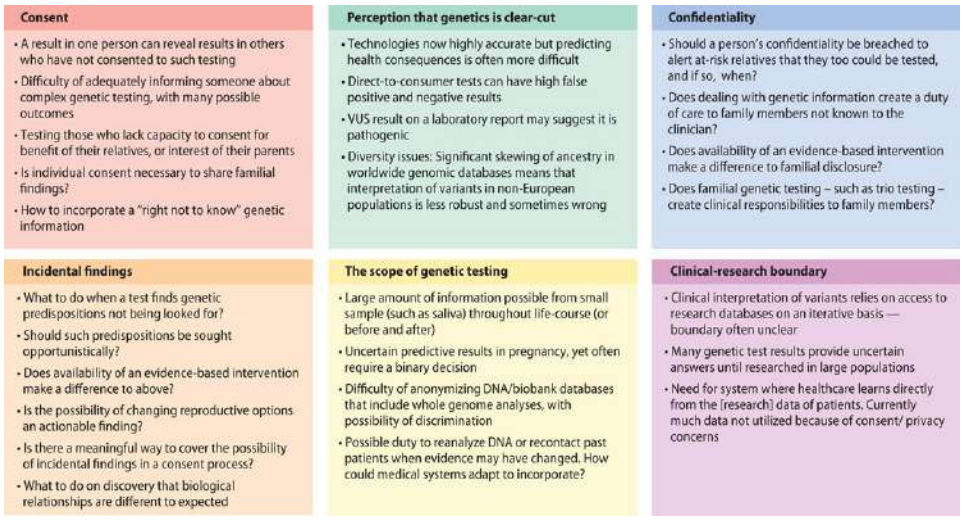


Figure 11.20 Some ethical issues in genetic testing. VUS, variant of uncertain clinical significance.

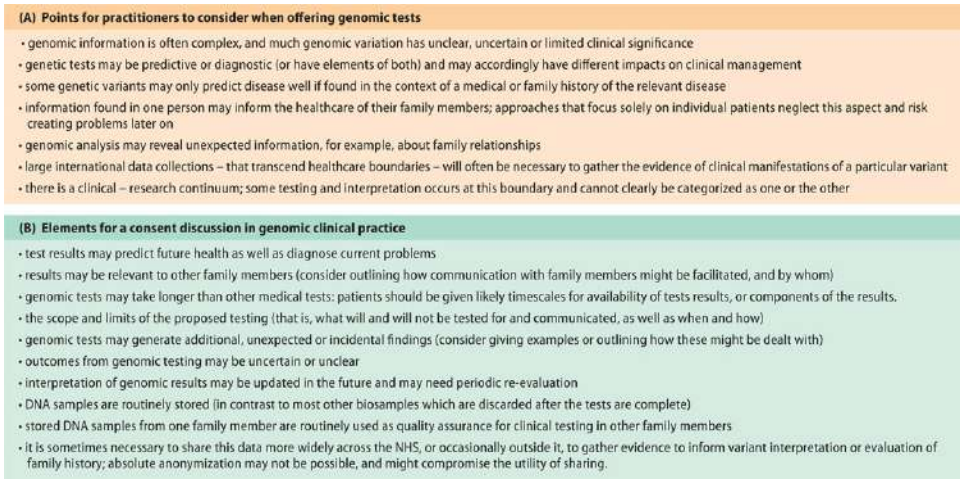
These issues arise to some extent in any genetic testing, but given that the diagnostic and predictive power of genetic testing in multifactorial disorders is weaker they are less stark in common conditions.

Genetic health professionals may find themselves in the rather uncomfortable position of meeting, and having access to information on, different family members who do not know about each other. Through genetic testing they may also discover that the biological relationships between some family members are different from the assumed relationships (as a result of misattributed paternity, unsuspected adoption, or sperm

donation—a case study will be presented below, in the section that includes incidental findings). Routine use of *trio testing*—in which samples of both parents' DNA determine whether suspect DNA variation in the child is *de novo* or not—significantly increases the chances of discovering misattributed biological relationships.

Consent issues in genetic testing

Consent to healthcare testing—including genetic testing—is an important aspect of respect for a person's autonomy. The consent process is meant to ensure that a person understands the nature and purpose of giving a sample, or of undergoing the medical intervention—see **Figure 11.21** for the aspects to discuss during consent for genetic testing, as recommended by the UK's Joint Committee on Genomic Medicine, and see **Figure 11.22** for a record of discussion form for clinical genetic or genomic testing proposed by the same organization.



(A) Points for practitioners to consider when offering genomic tests

- genomic information is often complex, and much genomic variation has unclear, uncertain or limited clinical significance
- genetic tests may be predictive or diagnostic (or have elements of both) and may accordingly have different impacts on clinical management
- some genetic variants may only predict disease well if found in the context of a medical or family history of the relevant disease
- information found in one person may inform the healthcare of their family members; approaches that focus solely on individual patients neglect this aspect and risk creating problems later on
- genomic analysis may reveal unexpected information, for example, about family relationships
- large international data collections – that transcend healthcare boundaries – will often be necessary to gather the evidence of clinical manifestations of a particular variant
- there is a clinical – research continuum; some testing and interpretation occurs at this boundary and cannot clearly be categorized as one or the other

(B) Elements for a consent discussion in genomic clinical practice

- test results may predict future health as well as diagnose current problems
- results may be relevant to other family members (consider outlining how communication with family members might be facilitated, and by whom)
- genomic tests may take longer than other medical tests: patients should be given likely timescales for availability of tests results, or components of the results.
- the scope and limits of the proposed testing (that is, what will and will not be tested for and communicated, as well as when and how)
- genomic tests may generate additional, unexpected or incidental findings (consider giving examples or outlining how these might be dealt with)
- outcomes from genomic testing may be uncertain or unclear
- interpretation of genomic results may be updated in the future and may need periodic re-evaluation
- DNA samples are routinely stored (in contrast to most other biosamples which are discarded after the tests are complete)
- stored DNA samples from one family member are routinely used as quality assurance for clinical testing in other family members
- it is sometimes necessary to share this data more widely across the NHS, or occasionally outside it, to gather evidence to inform variant interpretation or evaluation of family history; absolute anonymization may not be possible, and might compromise the utility of sharing.

Figure 11.21 Recommendations for the consent process from the UK's Joint Committee on Genomic Medicine.

RECORD OF DISCUSSIONS regarding testing and/or storage of genetic material

I have discussed genomic/genetic testing with my health professional and I understand that:

Family implications

1. The results of my test *may* have implications for other members of my family. I acknowledge that my results may sometimes be used to inform the appropriate healthcare of others. This could be done in discussion with me, or in such a way that I am not personally identified in this process.

Uncertainty

2. The results of my test *may* reveal genetic variation whose significance is not yet known. Deciding whether such variation is significant may require sharing of information about me including (inter)national comparisons with variation in others. I acknowledge that interpretation of my results may change over time as such evidence is gathered.

Unexpected information

3. The results of my test *may* reveal a chance of a disease in the future, and nothing to do with why I am having this test. This may be found by chance, while focusing on the reason for my test, and I may then need further tests to understand what this means for me. If these additional findings are to be looked for, I will be given more information about this.

DNA storage

4. Normal laboratory practice is to store the DNA extracted from my sample even after the current testing is complete. My sample might be used as a 'quality control' for other testing, for example, that of family members.

Data storage

5. Data from my test will be stored to allow for possible future interpretations.

Health records

6. Results from my test and my test report will be part of my patient health record.

Note of other specific issues discussed (*eg referral to particular research programmes, insurance*):

I agree to genetic/genomic investigations*	**DATE** _____/_____/_____
---------- **Patient/parent signature**	**Discussion undertaken by: (clinician's name and signature)**
----------	----------

Affix sticky label or fill in details

Patient name: _____

Date of birth _____/_____/_____

Patient address:_____

Genetics ref. _____ 1 COPY for notes, 1 COPY for patient to retain

**insert details here, eg to investigate the cause of my child's developmental delay / family history of cancer / heart disease etc*

Figure 11.22 Record of Discussion form following the consenting process, as recommended by the UK's Joint Committee on Genomic Medicine.

As a rule, the process of seeking consent ensures that a person understands the nature and purpose of the procedure, or intervention in question, thereby asserting their right to self-determination. This therefore

applies to individuals who have capacity (or competence) and is not possible for children or adults who lack capacity.

There are three essential criteria for legally valid consent:

1. The person providing consent must have sufficient, appropriate information to be able to make a decision.
2. They must be competent to make a decision.
3. The decision must be voluntarily given (free from coercion).

Qualifiers around consent can confuse this definition. For example, consent is not consent unless it is informed, so it is not always clear what the mantra of "informed consent" means. It might be more helpfully understood as "information that needs to be provided in order for a person to give their consent".

Verifying past consent, or updating consent, so that family members can benefit, may not be possible because contact may have been lost, or it may not be clinically appropriate because the family member seeking information may be concerned about his or her confidentiality being compromised. For example, a pregnant woman who wishes to undergo prenatal diagnosis may not want anyone to know about the pregnancy until the test results are available. The necessity to seek consent from another family member for release of the information could lead to a breach of confidentiality for the pregnant woman.

In much of medical practice, an intervention or treatment may be proposed because it will directly address a health problem or its treatment. Sometimes it is so integral to a patient's care that their consent might be assumed or inferred. For example, it would be unusual to seek specific, separate consent to check a person's blood pressure or cholesterol level. Inferring consent to a genetic or genomic test may be more problematic because such a test may reveal many different types of results, for example, current or future health problems; predispositions to conditions that may never manifest; information that requires more research before it is of clinical utility; or information that is of (greater) relevance to relatives.

Consent to genomic testing therefore needs to incorporate the complexity, uncertainty, and open-endedness of these many types of "results".

For adults who lack capacity, and are therefore unable to provide consent, a genetic test can be undertaken if it is believed to be in the best interests of the adult concerned. Those close to the adult, for example someone appointed as power of attorney for health and welfare, might help to inform what is in their best interest. It is important to remember, however, that capacity is decision-specific. A person may lack the capacity for certain decisions, but not others. For very young children (as in newborn screening) a person with parental responsibility may give consent for genetic testing. Generally speaking, adults are presumed to have capacity to consent, but may not have, whilst children (under 16–18 in most countries) are presumed not to have capacity but may demonstrate it depending on the age and maturity of the child.

A different consent lens for genomic testing?

Consent for medical investigations, including blood tests, is traditionally anchored at one point in time. It may need to evolve in the case of genomics, however, so that broad consent is obtained for the creation of a personal genomic resource that might be obtained, for example, at birth; subsequent clinical encounters seeking consent would be required to interrogate that resource with different questions at different stages of a person's life.

On a practical level, returning to test this resource multiple times (rather than repeating a blood test) will require a significant change in data storage and access within a health service and need to address as well as medical record-keeping so that, for example, a whole genome sequence can be accessed two decades after birth to answer questions about adult-onset predispositions. Such a new approach would also need to be able to revise data interpretations as new knowledge is brought to bear on them. Traditional notions of informed consent are difficult to apply to situations

where the possible outcomes are so unknown, both by virtue of the individuality of the genomic data, but also due to the complexity of navigating through that data to a "result".

The generation of genetic data is outstripping the ability to provide clinical interpretation

The threshold for initiating genetic testing as an investigation is being lowered. It may now be done simply out of interest (for example, ancestry testing) rather than through suspicion of a particular heritable factor. Consequently, the risk of overinterpreting genetic variation to predict disease increases. As mentioned in Section 11.4, direct to consumer (DTC) genetic tests are sold as providing answers; people buying them may understandably expect their results to be clearly predictive of future health.

One common pitfall is to compare a genetic test result to a "zero risk," rather than to a population risk. For example, after his polygenic risk score identified a 15 % risk of developing prostate cancer by age 75, a former British Health Minister declared that having a genetic test "may have saved my life". Experts immediately disputed the usefulness of this result, and his interpretation of it: in the UK the average lifetime risk of a man developing prostate cancer is 18 %. More men die *with* prostate cancer rather than *from* prostate cancer.

Careful framing of results (for example, comparing them with population risks) may mitigate the risk of over-interpretation. However, this relies on information being provided in an accessible manner. Users need to know how important it is to read the information carefully, which may not be obvious in the context of a societal discourse that tends to present genetic results as strongly predictive. The assumption that DTC genetic testing empowers people to reduce their future disease risk is undermined by evidence suggesting that learning about genetic predisposition to particular diseases rarely leads to sustained lifestyle change.

The ability to generate genomic data has substantially outstripped our ability to interpret its significance for an individual (see **Figure 11.23** for an

analogy), and while improvements in genomic technology are in many cases driving improvements in healthcare, interpretation of what such data means in a clinical setting, and what sort of intervention should be offered as a result, lags behind. The Global Alliance for Genomics and Health (GA4GH) predicts that by 2025, over 60 million people will have had their genome sequenced in a health-care context, but as suggested above, pathways for managing the output from genome sequencing are still in their infancy.

Figure 11.23 Improvements in genomic technologies can be likened to improved efficiency in the fishing industry. Single-gene approaches are like fishing for a particular fish, that one wants for dinner and knows how to cook. Whole genome approaches, by contrast, are like trawling the ocean bed. In such a case, one may not want to use the entire yield in one go, or be able to use it, or even to recognize what is in the net. One might often be better off throwing some of the catch back into the ocean to pick them out when they are matured. (Left, From Just Dance/Shutterstock.com. Right, From Susi Nodding/Shutterstock.com)

The detailed but unfocused approach of genomic tests gives opportunities to answer questions that go beyond the problems that led to a patient having a test (see incidental findings) but how well such questions can really be answered, and at what cost, is as yet unclear. At any given time, deciding which of the multitude of possible outputs from genomic tests should be considered a "result"—anticipated or otherwise—is challenging, not least because the links between many genetic variants and diseases are often unproven or poorly understood. Multidisciplinary input and collaboration will be key to interpreting the significance of genomic results.

New disease gene discovery and changing concepts of diagnosis

Exome and genome sequencing are powerful diagnostic tools. Take, for example, the Deciphering Developmental Disorders Project in the UK. It recruited patients with severe undiagnosed disorders (who generally already had had any currently available diagnostic genetic testing). Thereafter, exome sequencing was carried out in family trios (two parents plus affected child) to achieve a 40 % diagnosis rate for the first one thousand or so family trios in the study.

The search for a diagnosis has often been described as a journey, with parents of children with rare genetic disorders anticipating that a diagnosis may guide treatment, prognosis, acceptance and social support. However, identification of new rare disease genes may be changing the impact of receiving a diagnosis, and in many cases very little is known about the long-term effects of newly identified genetic conditions. Health professionals may find themselves in the position of learning about the effects of possible disease-causing variation(s) in a gene through meeting the patients in whom such genetic changes have been discovered.

Often the pathogenic variant will be in a gene newly thought to be linked to developmental disorders; there will be little, if any, published literature to draw on. Health professionals then have to speculate on whether the detected genetic change is the cause of the patient's health problems, and how it will impact the patient or their family in the longer term. This has

often led to patient support and awareness groups taking on an increasingly important role, as families gather to share their lived experience of newly diagnosed rare genetic conditions, in turn informing clinical services.

The agnostic approach of exome and genome sequencing is also challenging our previous concepts of existing genetic diagnoses. Genome-wide trawls often find apparently pathogenic variants in well-described disease genes in patients whose clinical phenotypes fall outside the boundaries of the phenotype expected. For example, loss-of-function variants in *SOX2* are known to cause anophthalmia and microphthalmia in addition to other phenotypes such as developmental delay and structural brain anomalies. Eye abnormalities were thought to be a key feature of *SOX2*-related disorders, and so *SOX2* would be requested as a genetic test only in patients who had absent or small eyes. Recently, via "genotype-first" approaches, loss-of-function *SOX2* variants have been found in people with developmental delay but without anophthalmia or microphthalmia, broadening the phenotypic spectrum associated with this gene.

Complications in diagnosing mitochondrial disease

Until recently, as noted above, a diagnosis of mitochondrial disease was often made rather late in a patient's investigative journey. With the advent of WGS plus the ability to accurately determine levels of mtDNA variants, such diagnoses might be made well before a patient exhibits the hitherto classical symptoms of mitochondrial disease. But earlier diagnosis may also mean that the range and severity of subsequent disease is difficult to predict. One reason for this is that mutant mtDNA level in blood invariably underestimates the levels present in less accessible, clinically manifesting, post-mitotic tissues, such as the brain.

Take the example of the m.3243A>G mutation in *MT-TL1*. It causes a relatively mild phenotypes (diabetes and deafness) at low mutant levels. But at higher levels, it causes complex disease presentations, including MELAS (mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes). As part of mainstreaming genetics and genomics, a diabetologist may

suggest that a young person with familial diabetes should undergo WGS testing, with the application of a monogenic diabetes virtual gene panel.

Identification of a m.3243A>G variant would not only constitute a primary finding and provide a diagnosis, but also imply a risk of developing additional future phenotypes. For some of these—such as hearing loss, cardiac involvement, and renal dysfunction—screening may alter the course of the disease. But for others, such as stroke-like episodes (SLEs), there is as yet no early intervention to alter the clinical course of the disease. As mentioned, the predictive value of extrapolating m.3243A>G levels in blood to brain tissue is limited, and this raises important questions about when and how to communicate uncertain findings, especially where there is no clinical action to offer as a result.

mtDNA single-nucleotide variants are passed on by females to their offspring. However, as a random small sample of wild-type and mutant mtDNA are "bottlenecked" into each individual ovum (as illustrated in Figure 7.17 on page 214), the resultant mutant load in the child can be considerably different to that in its mother, including a much more severe phenotype in the child. A woman found to carry the m.3243A>G variant may choose to proceed with a pregnancy, have prenatal testing with the option of terminating a pregnancy with high variant levels, or (in some jurisdictions) undergo mitochondrial donation *in vitro* fertilization (see below). Each option may result in considerable anxiety, and the woman may already have children who have therefore been *de facto* tested for what is typically an adult-onset condition.

Complications arising from incidental, additional, secondary, or unexpected information

The potential for discovering "other" information depends on several factors. It might depend on what question led to genetic testing (was it a diagnostic or screening test, for example), how broad or targeted the genomic analysis is, and what level of interrogation of the genomic sequence takes place. For example, if someone has a genomic test to

investigate a familial tendency to cardiomyopathy, finding a disease-associated *BRCA1* variant may be entirely unexpected information, and incidental to the question at hand. Whether this is possible will depend on the filters applied to the bioinformatic pipeline.

When a test is requested for a particular reason, there has been much recent discussion about how far findings that may indicate future disease should be routinely sought. Such findings are usually termed additional or secondary findings when a routine search for them has been done. Those discovered whilst looking for something else are reported as incidental, or unexpected findings. A gene panel approach looking for recessive conditions in childhood may, for example, find heterozygous gene variants conferring adult-onset cancer or neurological conditions.

The ability to find genetic variants unrelated to the clinical problem that a patient presents with are an inevitable consequence of the increased sensitivity of genomic testing. This is of course not so different from other types of clinical tests: a whole-body MRI scan done to investigate one symptom may reveal a quite unsuspected tumor or aortic aneurysm, for example. But there are at least two subtle differences for incidental findings in genetic tests, as listed below. First, they may predict clinical manifestations many decades from the point of their discovery. Secondly, they may also predict clinical manifestations for close relatives.

Opportunistic finding of (other) health risks could be considered helpful, of course; but working out how to handle this information raises difficult questions. In 2013, the American College of Medical Genetics and Genomics (ACMG) suggested that when performing clinical sequencing, laboratories should automatically seek and report pathogenic variants in 56 genes associated with “medically actionable” conditions (revised in 2021 to 73 genes). The main rationale was the potential benefit of diagnosing disorders where preventative measures and/or treatments were available, with the aim of improving health.

The ACMG recommendations above proved controversial. The debate centered around whether patients should have a right to choose not to know such information. Other questions are yet to be fully addressed. They

include the following: (a) What constitutes a "medically actionable" finding? (b) What is the predictive value of such findings in the absence of a phenotype or family history of the relevant disorder? and (c) How do we reconcile this with the statement that this search is not validated for population screening?

Analysis of data from the 1000 Genomes cohort demonstrated that approximately 1 % of "healthy" people will have a "medically actionable" finding in one of the 56 ACMG-listed genes. What this might mean on an individual basis, however, is often unclear. Most of our knowledge of the effects of variation in gene *X* has been gathered by studying people identified as having a gene *X* variant, and they have been tested because of a personal or family history of gene-*X*-associated disease. That inevitably biases the sample from which our conclusions are drawn.

It is less clear what it might mean to find, for example, an apparently pathogenic variant in a gene linked to cardiomyopathy in a person with no personal or family history of heart problems (see **Box 11.7** for a case history). This has important implications for cascade testing of relatives. To what extent should testing and subsequent screening be offered in a family based on an incidental finding of a genetic variant thought to be predictive of a particular condition, if there is no clinical evidence that anyone in the family, including the person in whom the genetic variant in question was first identified, is actually affected by it?

CLINICAL BOX 19 CASE STUDY: POOR PREDICTIVE VALUE GENETIC TESTING IN ABSENCE OF CLEAR CLINICAL PHENOTYPE

A two-year-old boy was investigated for "absence spells". He had no loss of consciousness, and after being investigated in detail for epilepsy, no abnormalities were found. The community pediatrician attempted to reassure the parents that this is a normal feature in some children, and that he would likely grow out of these spells. As a precaution, however, the boy was referred to a pediatric cardiologist, who also found no

abnormalities: his baseline ECG was defined as within normal limits, and he had no family history (to 3rd-degree relatives) of any cardiac problems.

The cardiologist had been to a presentation about mainstreaming genetics and realized that long QT syndrome (leading to increased chance of sudden cardiac death) can be difficult to diagnose in childhood. He therefore requested screening of a gene panel "to exclude long QT syndrome". A *KCNQ1* variant associated with long QT was identified, and described on the laboratory report as "likely to be pathogenic". A reveal device was inserted but no abnormalities in the boy's QT interval were recorded during subsequent absence spells. As a precaution he was treated with beta blockers, and cascade genetic testing of his family was initiated. This revealed that his three-year-old sister, father, paternal aunt (and her two children, aged 4 and 8) and paternal grandfather all carried the same variant. Cardiac investigations of their phenotype, at rest, with exercise, and pharmacological challenge were all normal or equivocal. All carriers in the family were prescribed beta blockade and two members of the family were referred for possible implantable cardiac defibrillator insertion.

LEARNING POINTS

The significance of genomic variants found in the absence of a phenotype can be very unclear. It is easy to see why investigations and treatments were requested as a precaution here, but also quite possible that a significant health resource has gone into investigating and treating many members of a family when no-one is at increased risk of sudden cardiac death. Such psychological and financial costs are significant, and have the potential to be burdensome to mainstreaming agendas. It is important that the data from these sorts of examples are collected systematically, and learned from, to improve future practice.

Broad genomic testing also has the potential to detect carrier status for recessive and X-linked conditions. On a disorder-by-disorder basis, being a

carrier for a genetic condition is very rare (with notable exceptions such as hemochromatosis and cystic fibrosis). But it is very common, and "normal", to be a carrier for a genetic condition. A population study simultaneously testing carrier status for 100 or so recessive disorders in nearly 25 000 people found that 25 % were carriers for at least one of the disorders, and 5 % were carriers for multiple disorders.

For most people, being a carrier will have no impact on their life at all. However, if their partner happens to be a carrier for the same condition, the implications can be profound: each of their children would have a 25 % chance of being affected by the genetic condition. This is particularly relevant for couples known to be biologically related, and couples with common ancestry (who will have a higher chance of both being carriers for the same recessive condition).

Because of the increased scope of carrier screening and because being a carrier for one or more recessive genetic conditions is very common, carrier results for recessive genetic conditions are increasingly conveyed on a couple basis. That is, carrier status is only communicated if relevant in the context of a particular relationship, where both individuals in a couple are carriers for the same condition (see Section 11.4 above). Making the status of "being a carrier" part of normal variation would be a welcome development, but the notion of a couple's result rather than an individual result needs careful consideration, not least in terms of recording this information in medical records.

Interestingly in the UK's 100 000 genomes project, although participants were offered a subset of variants on the ACMG list as "additional, looked for findings" no additional findings have yet been communicated from the project (at the time of writing). That is so although consent for such extra investigations was obtained up to six years previously; many research participants no longer recall what they consented to. What is urgently needed are long term implementation projects that assess the penetrance of these variants in a general population as well as an evidence basis for interventions offered (see **Box 11.8** for a case history)

Consent issues in testing children

Genetic testing of children raises additional consent issues. Should children at risk be tested for adult-onset conditions? Or screened for carrier status for serious recessive disorders? The answer to both questions is usually no; unless there is clear medical benefit at that time, testing should be delayed until the child has the capacity to make the choice.

Sometimes parents will request such testing because they consider that such knowledge would be helpful. However, if there is no chance of a childhood onset of the condition, and no interventions or actions that can be taken now to alter the course of the condition, then a plethora of international guidelines recommend deferring such testing until the child is old enough to consent themselves. This also respects the child's "right to an open future" whereby decisions that can be delayed are, so that options for the child are not curtailed. This applies especially to conditions in which adults might sometimes choose not to be tested; testing during childhood would then deny the child the right not to give consent that he or she could exercise as an adult.

Current professional guidelines in many countries therefore stipulate that children should not normally be tested in this way unless there is clear medical benefit in early testing. Testing for familial hypercholesterolemia is one such example: early detection of a pathogenic *LDLR* mutation offers the possibility of prevention by lowering LDL-cholesterol through dietary changes and medication, and testing in *LDLR* mutations families is recommended from the age of 10 years.

CLINICAL BOX 20 CASE STUDY: MISATTRIBUTED GENETIC PARENTAGE AS AN EXAMPLE OF AN INCIDENTAL FINDING

Meena and Joe are seen in the genetics clinic after their daughter Ana is born with serious health problems. Whole exome sequencing (WES) on a sample from Ana finds two pathogenic variants in a gene associated with a

severe autosomal recessive condition. Further testing is needed to ensure that these variants were inherited on separate chromosomes, one from each parent. If so, Ann has no working copy of the gene; the true cause of her health problems has been found. Tests on parental samples show that Meena has one of the variants; Joe has neither. Further testing shows that Joe is not the biological parent of Ana.

Meena and Joe had previously been told that, as a couple, the chance of their future babies being affected by the condition was likely to be 1 in 4 (25 %). They were told that if a genetic cause for Ana's health problems were found, they could have prenatal genetic testing in future pregnancies. However, as Joe is not the genetic parent of Ana, the chance of Meena and Joe having a baby with the autosomal recessive condition would be very low. Prenatal genetic testing, with its associated miscarriage risk, would not be indicated.

KEY POINTS

- This scenario is most likely to be one of misattributed genetic paternity. However, gamete donation may result in other misattributed genetic relationships.
- Genetic testing can reveal unexpected social information as well as medical information; ideally this possibility should be made clear during the consent process, although the presence or absence of consent will not necessarily help to determine whether, when or how such a finding should be disclosed.
- While a clinician may feel uncomfortable introducing this type of "social" information into discussions, it can have medical relevance, for example, in predicting recurrence risk for future pregnancies.
- Different professional duties may arise when responding to existing information than considering whether potential information should be sought.

- Trio testing (analyzing the genome of both parents together with the child under investigation to improve the diagnostic yield) is now standard practice in the investigation of rare diseases.
- The interpretation of variation discovered through WES or WGS (whole genome sequencing) is challenging; there is an enormous amount of variation across the genome even within genes that can cause severe diseases, so that demonstrating inheritance from both parents (or in the case of a dominant condition demonstrating the finding in Ann is *de novo* and not inherited from a healthy parent) is important for diagnostic interpretation.
- The possibility of discovering such findings increases with trio testing. Some services have a policy of not disclosing such findings and/or labelling a sample as "failed" if it is not genetically related to the child. However, this may lead to repeat sampling if the reasons for the trio failure is not made explicit.

Ethical and societal issues in prenatal diagnosis and testing

Prenatal diagnosis for serious genetic disorders has long been available in many developed societies. As noninvasive prenatal screening technology develops and is standardized, the miscarriage risk of invasive procedures (chorion biopsy or amniocentesis) is becoming less of an obstacle to prenatal diagnosis; it also allows for a diagnosis at an early stage of pregnancy. After a diagnosis of a harmful genetic variant, terminating a pregnancy is accepted in many societies, although support is often far from universal. First trimester, or early second trimester, terminations are usually less traumatic medically and socially for all involved.

For those who consider that terminating a pregnancy can be justified for serious genetic disorders, another issue remains: where do we draw the line that divides serious disorders from non-serious disorders? Some couples might wish to contemplate termination for what many other people might consider mild disorders, such as congenital deafness.

And as we move towards being able to routinely analyze a fetal genome through non-invasive tests, further questions arise. In the past, prenatal genetic testing was usually only offered when a particular fetal phenotype was noted or suspected, meaning that filtering and interpretation of genetic variants identified could be anchored in attempts to explain an existing health concern. Advanced genomic testing is now increasingly used in pregnancies where there is no prior suspicion of genetic abnormality, producing information on genotype without the phenotypic data required to give it meaning. This increases the difficulty in predicting whether, and how, particular genetic variants might affect future development and health (see page 464 for a worked example on prior probability).

A challenge to healthcare scientists, clinicians, and parents, therefore, is deciding what qualities prenatal genotypic variation should have in order to be constructed as a "result". At the same time, such tests are often re-requested in order to make binary decisions about whether to continue a pregnancy or not. A range of professional organizations are developing guidelines on the use of advanced genomic testing during pregnancy. However, the discovery of ambiguous findings—such as variants with uncertain clinical significance, susceptibility loci for neurodevelopmental problems, and susceptibility to adult-onset diseases—remains a difficult management problem. Any decision to terminate—or not—will need to be made well before we know whether any of these will manifest.

Preimplantation genetic diagnosis (PGD)

For any single gene disorder, a proportion of embryos produced will be unaffected: 50 % in the case of autosomal dominant transmission where one parent carries the pathogenic variant, and 75 % in the case of autosomal recessive condition (although two-thirds of the unaffected group will be carriers). After genetic testing of embryos, PGD offers the option of implanting just those embryos expected to have a normal phenotype from

genetic testing (see Figure 11.18 on page 460 for the practicalities). It thus avoids the difficult choice to terminate a pregnancy.

In the UK, any PGD is tightly regulated by the Human Fertilisation and Embryology Authority (HFEA), and practitioners are asked to consider the welfare of the future child in their decisions around which embryos to implant. Difficult questions can arise concerning the purpose of PGD, and are hotly debated. Instead of using PGD to avoid transmission of a genetic condition, for example, might it be used instead, to select an embryo purely on the basis of a predicted HLA antigen profile that is a close match to that of an existing child in the family who needs a tissue transplant?

Newborn genome sequencing

In the preface to the previous edition of this book, although the prospect of whole genome screening of newborns (neonates) might have seemed on the distant horizon then, in 2014, we did ask this question: *might we soon live in societies in which genome sequencing of citizens becomes the norm?* Well, the time when this happens is looking much closer now. In the Americas and Europe, extrapolations of recent genome sequencing suggest that by 2030, at least 60 million citizens will have had their genomes sequenced. And should the goals be met of China's 15-year Precision Medicine Initiative—at $9.2 billion, the largest of its kind—another 100 million genomes may be delivered by 2030.

In the UK, following the 100 000 genome project, there are plans to analyze the genomes of five million people (Our Future Health—https://ourfuturehealth.org.uk/). And the vision to offer newborn screening via whole genome sequencing (**Figure 11.24**) is much closer to a realistic endeavor.

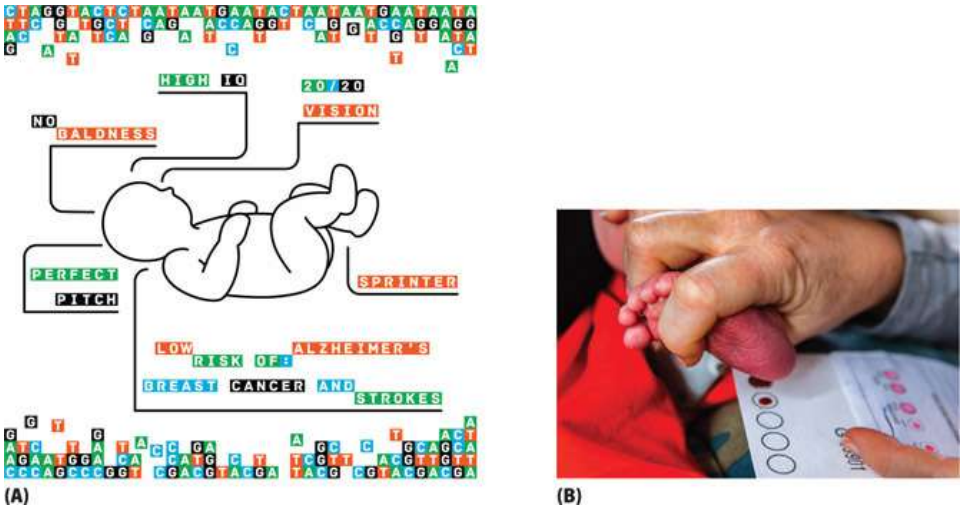


Figure 11.24 Will genome sequencing of neonates be the future norm?(A) Prenatal whole genome sequencing. From Olga Boat / Shutterstock.com, permission. **(B)** From Valmedia / Science Photo Library, with permission.

At the time of writing, the UK government has just published the results of a public consultation exercise on the idea of newborn screening by whole genome sequencing (WGS), available at: https://www.gov.uk/government/news/public-dialogue-on-the-use-of-whole-genome-sequencing-in-newborn-screening. Its conclusions were as follows:

1. It would be acceptable to identify a wider set of conditions than the current newborn screening program if they impact the infant in early childhood and there are treatments and interventions to cure, prevent, slow progression, or personalize treatments.
2. A comprehensive genetic database should be established so that people from ethnic minority backgrounds are not disadvantaged by receiving more uncertain, or less accurate, diagnoses than the rest of the population.
3. The full complexities of whole genome sequencing must be recognized within any consent processes including:

    a. its implications for the wider family
    b. that 21st-century families come in many forms

c. while parents give consent on behalf of the newborn, the child may have a different view as they grow up, including on whether their genomic data are used for research
d. that the screening test has potential to look for many more conditions than current newborn screening tests, and that some of these may not appear for many years, or be poorly predicted by genetic variation alone.

The dialogue participants confirmed that in many ways, sequencing and analyzing genomes is the easy part. The really difficult questions revolve around how predictive the results are, what conditions it would be acceptable to look for, what information to give to whom and when, and how to help parents make informed choices about tests that could have important implications for their child, for themselves, and maybe for others in their family over many years.

Ethical and social issues in some emerging treatments for genetic disorders

Rapid developments in diagnosing and delineating molecular disease mechanisms have advanced treatment prospects for an increasing number of genetic conditions. They include the use of therapeutic monoclonal antibodies and other proteins produced by genetic engineering (Section 9.2), and various gene and RNA therapies (Sections 9.3 and 9.4).

The vast majority of gene therapies involve genetic modification of the somatic cells of patients, and have no consequences for future generations. Germline gene therapy has potential consequences for future generations and is widely banned. However, a recent proven treatment for certain severe mitochondrial diseases is effectively a type of germline gene therapy in which donor mtDNA becomes incorporated into the germ line. That happens by a type of *in vitro* fertilization in which mitochondria in the early embryo or egg cell are replaced by mitochondria from an oocyte donor (see

Figure 9.26 on page 356.). This type of therapy is legally permissible in the UK.

If the technology of genetic modification using CRISPR-Cas or similar genome editing method advances in the future so that it becomes highly efficient, and safe to carry out, the prospect is raised of germline genome editing. That may open the door to *genetic enhancement*, the prospect of modifying the genome to select for some quality perceived to be desirable.

We give an overview of some of the ethical and social issues raised by treatments for genetic disease and genetic enhancement in **Table 11.10** and enlarge on two areas immediately below this: inequality of treatment provision; and treatment for mitochondrial disease by mitochondrial replacement. And we conclude with a section on the prospects and ethics of germline modification of nuclear DNA and genetic enhancement.

TABLE 11.10 SOME ETHICAL AND SOCIAL ISSUES IN TREATING GENETIC DISEASE AND GERMLINE GENOME MODIFICATION

Treatment/genome modification	Ethical and/or social issues
Drugs provided by the pharmaceutical industry	Inequality of availability because of moderate to sometimes large costs
Invasive procedures required by new treatments	Concerns about invasive procedures needed in tiny infants, such as intubation and tracheostomy to administer Nusinersen in infants <1 year of age with severe spinal muscular atrophy. Invasive treatment can be hard to stop, once begun, even when futility becomes clear

mAbs, monoclonal antibodies.

Treatment/genome modification	Ethical and/or social issues
Genetically engineered mAbs and other "recombinant" proteins	Gross inequality of availability, as a result of huge annual manufacture expenses
Licensed somatic gene therapy and RNA therapeutics	Gross inequality of availability through huge expense of treatment
Mitochondrial replacement therapy	Some ethical concerns about germline alteration, even although the donated mtDNA is natural, as opposed to artificially altered
Germline genomic editing	Major ethical concerns about alteration of the germ line having unforeseen consequences for future generations

mAbs, monoclonal antibodies.

Inequality of treatment availability

The press typically reports advances in treatment of genetic disorders with great enthusiasm. There is often little mention of the downsides, which include the huge inequality of availability of many of these treatments because of costs that can sometimes be staggering. That can also apply to treatments used for decades, allowing refinement of the production process, not just major advances that have recently burst on the scene that might be expected to be initially expensive.

Take the example of hemophilia. An estimated 20 000 people in the US are living with this inherited bleeding disorder, and more than 60 % of them have moderate or severe hemophilia requiring lifelong treatment with expensive drugs and clotting factors. During the 1960s the average life

expectancy for a patient with hemophilia was ~12 years. Recombinant factor VIII was made in 1984 and approved for medical use in the US in 1992; now people diagnosed with hemophilia can anticipate a near-normal life expectancy if treated with recombinant factors VIII or IX. In 2020 the *American Society of Hematology Clinical News* reported that treating an adult patient by replacing factor VIII or IX with the genetically engineered recombinant protein, or by using the newer bispecific antibody emicizumab, costs somewhere in the region of $300 000 to $500 000 *per year*.

Of course, experimental gene therapies and RNA therapeutics are also very expensive; few have been licensed thus far. Conventional drugs produced by the pharmaceutical industry to treat or prevent genetic disease can vary in cost and availability. For example, statins and beta blockers are not so expensive and are widely available, but new drugs can be very expensive: lifetime treatment of cystic fibrosis using Vertex Pharmaceuticals' effective Trikafta drug (a combination of the elexacaftor, tezacaftor, and ivacaftor drugs described in Section 9.2) costs more than $6 million dollars per patient.

The ethics of treating mtDNA disorders by mitochondrial donation

Recall that mitochondrial DNA (mtDNA) is present in hundreds to thousands of copies per cell and is strictly maternally inherited. Disorders in which a person is homoplasmic (100 % mutant mtDNA), or has a high percentage of mutant mtDNA, can result in a clinically severe disorder (mitochondria are the batteries of a cell; defective mitochondria especially affect organs needing the most energy: brain, muscles, and heart). Disorders like these are incurable, and reproductive choices have been mainly limited to egg donation or preimplantation genetic diagnosis to select embryos with the lowest percentage of variant mtDNA.

A woman with a *heteroplasmic* disease-causing mtDNA variant has a mix of mutant and normal mtDNA. She might have very few symptoms or

be unaffected, but because of the *mitochondrial genetic bottleneck* (Figure 7.17 on page 214) only a very few of the available mtDNAs pass from early primordial germ cell precursors into the egg, but in an unpredictable fashion. As a result, a heteroplasmic woman might quite often produce eggs with a high load of mutant mtDNA.

Mitochondrial replacement therapies (also known as *mitochondrial donation*) were detailed at the end of Section 9.4. Two different *in vitro* fertilization methods can be used for this purpose (described in Figure 11.18 on page 460). The essential point is that in the case of a woman with a heteroplasmic mtDNA variant, the normal nuclear DNA present in the unfertilized (or fertilized) egg is removed, then injected into an enucleated donor egg containing healthy mitochondria (before being fertilized *in vitro*), or into the already fertilized enucleated donor egg. Fertilization occurs using sperm provided by the prospective father. The resulting "three-parent babies", as sensationally reported by the world's press, might give the erroneous impression of three equivalent genomes passed to the child; of course, the donor contributes a mitochondrial genome only, just 0.0005 % of the size of each of the two parental genomes.

The UK was the first country to regulate the use of this approach, after the Human Fertilisation and Embryology Authority (HFEA) conducted a scientific review and public consultation that informed a parliamentary debate to approve their use in a clinical setting. Interested readers can find a recent review of regulation of the method in different countries at PMID 31961722 and a neat summary of the ethical debate in a Nuffield Council of Bioethics report at https://www.nuffieldbioethics.org/assets/pdfs/Novel_techniques_for_the_prevention_of_mitochondrial_DNA_disorders.pdf.

The ethics of germline gene modification for gene therapy and genetic enhancement

Mitochondrial donation therapy, as described immediately above, may be a type of germline gene therapy, but it has a clear benefit with arguably few

ethical concerns. Consider that it is simply a question of replacing a small amount of damaged genetic material, and, importantly, the replaced genetic material has not been artificially edited or designed in any way. Instead, the procedures essentially involve replacing damaged mitochondria by healthy mitochondria containing natural mtDNA. The treatment is new and there has, as yet, been limited experience, but it will be important to carefully monitor the safety of these treatments.

Germline modification of the nuclear genome, by contrast, would involve making artificial genetic changes to germline cells that might be transmitted down the generations. The current genetic technologies for modifying the nuclear genome are led by CRISPR-Cas genome editing, but impressive though this technique might be, the technology is currently imperfect. When the CRISPR/Cas9 system was used by a group in China recently to correct pathogenic variants in the *HBB* and *G6PD* genes in human zygotes, the efficiency and accuracy of the correction procedure was variable. Errors, notably "off-target" effects can be introduced unknowingly that might have harmful consequences for generations to come.

Even if the technology were refined and matured to the stage where it was efficient on every occasion with no off-target effects, there might still be unintended consequences because of our imperfect knowledge of the complex nuclear genome. At this stage, we might step back and ask why we would ever want to carry out germline modification of the nuclear genome. Studying genetically modified germline cells in culture might be considered desirable for our basic understanding of these cells, but could germline genetic modification that may be transmitted through reproduction ever be ethically acceptable? We look at two scenarios below.

Germline (nuclear) gene therapy

A review published in *Nature* in 1998 reported that panelists at a symposium on "Engineering the human germ line" held in 1998 at the University of California almost unanimously argued in favor of

implementing germline gene therapy, once techniques for altering the germ line could be conducted safely and effectively in human embryos, and "regardless of the concern that its use might lead to an ethical morass" (PMID 9537311). James D Watson was reported as telling the symposium that "scientists should proceed unhindered towards germline engineering" and advocating that such therapy must be spared excessive regulation, adding: "if there is a terrible misuse and people are dying, then we can pass regulation". The European view was different: after a bioethics convention produced by the Council of Europe in 1997 a total of 22 European states supported the argument that such genetic manipulation should not be carried out if the aim was to introduce a permanent modification in the genome (that might be transferred down the germ line).

A possible argument in favor of germline gene therapy might be the desire to eliminate the risk of an inherited disease to future generations at the population level. However, for recessive conditions, only a very few of the disease alleles are carried by affected people (the great majority are in healthy heterozygotes), and most serious dominant or X-linked diseases are largely maintained in the population by recurrent mutation.

And, in a correspondence letter to *Nature* in 1998, in response to coverage of the "Engineering the human germ line" symposium, Anne Maclaren pointed out that there was simply no need for germline gene therapy (PMID 9565021). Rather than seek to "correct" harmful sequences in embryos, one could test cells taken from the early embryo using preimplantation genetic diagnosis, and then just implant the normal cells. After all, for nuclear genes, the highest risk for transmitting disease comes from variants associated with single-gene disorders, with at its highest, a 50 % risk in the case of dominant disorders, leaving 50 % normal embryos.

Genetic enhancement and "designer babies"

To some people, the enthusiasm for germline nuclear gene therapy might have concealed an ulterior motive. In her 1998 correspondence letter to

Nature, Anne Maclaren publicly asked James Watson whether he had simply forgotten about the possibility of preimplantation genetic diagnosis (which would make germ-line gene therapy pointless), or whether it was germline engineering for *genetic enhancement* that he wished to proceed unhindered. No reply was disclosed.

Caught up with this selection of desired traits through genetic enhancement is the prospect that people might wish to use *in vitro* fertilization and preimplantation diagnosis simply to detect and select embryos that offer certain desired qualities, and reject the rest even though they do not harbor a genetic condition.

A demand for "designer babies" with multiple desired qualities might conceivably become a reality in the future if we had a much higher level of information about which genes to modify and if genetic manipulations on the germ line were to be so efficient that the technology became extremely safe. Some people argue that there would be a moral imperative to undertake human germline editing once the techniques are sufficient advanced, but in the real world this is mitigated by the fact that it is not usually possible to ensure a better life.

The moral arguments above also tend to rely on an overly deterministic view of a genome sequence, and the role of variation within in it, in the etiology of the disease or traits. Certainly, most common diseases cannot simply be attributed to specific genetic variants that we could edit away. Multiple, poorly understood genetic and environmental factors interact to influence the expression of diseases with a genetic component, even well understood "monogenic" disorders. As mentioned above, population-level genome analyses are now demonstrating that many genetic "mutations" are much less predictive than previously thought. Furthermore, human genome editing might introduce new risks just as it reduces old ones; or remove protections not yet clearly delineated. Similarly, the genetic basis of character traits, or particular talents, is so complex and multifactorial that acting on any such moral imperative, even if this was uniformly agreed, remains the terrain of science fiction for the foreseeable future.

SUMMARY

- The analytical validity of a test evaluates how well the assay measures what it claims to measure. Many genetic tests with high analytical validity have low or absent clinical validity.
- A genetic test assay is said to have a high sensitivity if a high proportion of all people with the condition are correctly identified as such, and a high specificity if a high proportion of all people who do not have the condition are correctly identified as such.
- Healthcare service-led genetic tests may be used to confirm a clinical diagnosis, predict the likelihood of developing or transmitting a genetic disorder, predict the clinical course, or help monitor disease. (Some additional tests assay for drug responses, as described in Section 9.1.)
- Most genetic tests are designed to detect chromo-some abnormalities or pathogenic DNA variants. They may involve scanning for an undefined abnormal DNA variant, or testing for a one or more defined pathogenic variants.
- Rather than detect causative genetic variants, some genetic tests indirectly assay a convenient disease-associated characteristic, either altered expression products, altered gene function, or a characteristic disease biomarker such as an abnormally elevated metabolite.
- Genotyping specific single-nucleotide variants often makes use of pairs of allele-specific oligonucleotides that hybridize specifically to template DNA with either the normal or variant sequence.
- Targeted DNA sequencing means using biotin-streptavidin capture of desired genome sequences so that they can be selectively analyzed by DNA sequencing. It may be used to

capture multiple genes associated with a specific disease or group of diseases (gene panels), or a whole exome (containing all coding sequences and some untranslated sequences).

- Whole genome sequencing (WGS) may be used in identifying rare disease genes but is comparatively expensive. Identification of a pathogenic variant is not eased by the sheer number of background genetic variants, and difficulties in interpreting some variants.
- In broad genome scans additional pathogenic variants may incidentally be found that are associated with phenotypes other than those for which the test was ordered (incidental or secondary findings).
- As the costs of WGS fall, virtual gene panels are increas ingly used: bioinformatic filters are applied to screen out most of the genome sequence of patients, leaving genes of interest, such as all genes associated with heart disease or mitochondrial disease.
- Assessing the pathogenicity of sequence variants can be difficult. Identifying precedence (previous occur-rences of the variant), genomic constraint (strong evolutionary conservation of the sequence at the mutation site), and rarity of the sequence variant is often helpful.
- Genetic screening means carrying out proactive assays to identify individuals at increased risk of carrying harmful genetic variants. Past approaches to target particular communities or populations with an elevated incidence of, for example, a particular auto-somal recessive disorder, are rapidly being replaced by screening of entire populations.
- Traditional prenatal diagnosis has used invasive pro cedures to recover and analyze fetal cells from early pregnancy. In preimplantation diagnosis, genetic testing occurs on embryos

produced by *in vitro* fertilization in the context of assisted reproduction.

- In preconception screening, couples can be screened for many different autosomal recessive conditions. Carrier couples then have more reproductive options than when this is first discovered in pregnancy, or after birth.
- In noninvasive prenatal testing, samples of freely cir culating DNA recovered from maternal plasma are analyzed. The plasma DNA is a mixture of fetal and maternal DNA issuing from degraded cells. It can be analyzed to infer fetal DNA variants and the fetal genome sequence.
- Cascade testing means testing of relatives after iden tifying a person with a pathogenic mutation. Relatives will be at a higher risk (than the general population) of being carriers of a recessive disorder or chromosome translocation, or of developing a childhood-onset or late-onset dominant disorder.
- Pre-symptomatic diagnosis can be carried out on asymptomatic individuals at risk of developing a genetic disorder later in life. If a person is identified as carrying the mutant allele, follow-up screening can be carried out, and in some cases treatment regimes can be followed to reduce disease risk.
- In direct-to-consumer genetic testing, commercial companies carry out genetic tests and feed back results without involving healthcare professionals. The main focus may be on genetic ancestry, but predictions about future health risks are often given.
- Genetic testing for susceptibility to common diseases can identify individuals at increased disease risk; because the disease susceptibility here is multifactorial, even the best polygenic risk tests can measure just the genetic component to disease risk.

- Mainstreaming genetics envisages incorporation of genetic testing into mainstream medicine—much as, say, radiology or hematology has been incorporated in the past. We still need radiologists to interpret complex imaging; it seems likely that genetic professionals will be asked to fulfil similar advisory roles.
- Clinical genome sequencing can identify pathogenic variants, but if done without reference to a phenotype (for example prenatally), predictions are often much less clear than people expect. Furthermore, each person has a large number of variants whose clinical significance is weak or uncertain.
- Clinical genome sequencing is being incorporated into existing healthcare systems of many economically advanced societies. Significant bioinformatic and electronic networking challenges persist, as do ethical concerns about releasing data when we currently have imperfect knowledge of the clinical significance of many variants.
- Genetic testing is unusual in that the results often have potential implications for close relatives, as well as for the person tested. Professionals may sometimes need to balance preserving the confidences of one person with the prevention of harm to relatives (by alerting relatives to particular screening, for example).
- Clinical information about the person tested should be held in confidence, but the genetic factor that led to a diagnosis might be considered confidential to several family members. This means that relatives can sometimes be alerted to their risk without breaking the confidences of others.
- Consent for genetic testing should address the com plexities of genetics including implications for family members, the fact that uncertain information may be found, or that interpretation of findings may change over time.

- Genetic testing of children should only be done if it benefits them as children. If the test predicts adult-onset conditions for which there is no beneficial intervention in childhood, testing should usually be delayed until the child can be involved in the decision-making process. Disclosure of data that would enable such predictions to be made—from whole exome sequencing at birth, for example—should not be done until it benefits the child.
- Treatment of genetic disease normally has direct con sequences for just the person treated. Genetically modifying the germ line would potentially have consequences for future descendants and is widely banned.
- Mitochondrial replacement (also called donation), a way of avoiding transmission of severe mitochondrial disorders, is a type of germ line modification that simply replaces the mitochondria of a heteroplasmic woman by intact mitochondria from a donor egg cell.

QUESTIONS

Questions can be downloaded by visiting the following link, under Support Materials: www.routledge.com/9780367490812.

FURTHER READING

Genetic testing overviews and resources

Genetic Testing Registry [An electronic resource established by the US National Institutes of Health to serve as a central repository of genetic tests]. Available at: https://www.ncbi.nlm.nih.gov/gtr/

Katsanis SH & Katsanis N (2013) Molecular genetic testing and the future of clinical genomics. *Nat Rev Genet* 14:415–426; PMID 23681062.

Korf BR & Rehm HL (2013) New approaches to molecular diagnosis. *J Am Med Assoc* 309:1511–1521; PMID 23571590.

Identifying chromosome abnormalities and large-scale DNA changes

Cusenza VY (2021) Copy number variation and rearrangements assessment in cancer: comparison of droplet digital PCR with the current approaches. *Int J Mol Sci* 22, 4732; PMID 33946969.

Mann K & Ogilvie CM (2012) QF-PCR: application, overview and review of the literature. *Prenatal Diagn* 32:309–314; PMID 22467160.

Schaffer LG (2013) Microarray-based cytogenetics. In Gersen SL & Keagle MB (Eds), *The Principles of Clinical Cytogenetics*, 3rd ed., pp. 441–450. Springer.

Wapner RJ (2012) Chromosomal microarray versus karyotyping for prenatal diagnosis. *N Engl J Med* 367:2175–2184; PMID 23215555.

Willis AS (2012) Multiplex ligation-dependent probe amplification (MLPA) and *prenatal diagn*osis. Prenatal Diagn 32:315–320; PMID 22467161.

Genotyping point mutations and DNA methylation profiling

Heyn H & Esteller M (2012) DNA methylation profiling in the clinic: applications and challenges. *Nat Rev Genet* 13:679–692; PMID 22945394.

Kwok PY (2001) Methods for genotyping single nucleotide polymorphisms. *Annu Rev Genomics Hum Genet* 2:235–258; PMID 11701650.

Syvanen A-C (2001) Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat Rev Genet* 2:930–942; PMID 11733746.

von Kanel T & Huber AR (2013) DNA methylation analysis. *Swiss Med Wkly* 143:w13799; PMID 23740463.

Genome-wide and disease-targeted sequencing in mutation scanning

Choi M (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA* 106:19096–19101; PMID 19861545.

Rehm HL (2013) Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* 14:295–299; PMID 23478348.

Yang Y (2013) Clinical whole exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 369:1502–1511; PMID 24088041.

Interpreting and classifying sequence variants (see also Table 11.5 on page 447)

Findlay GM (2021) Linking genome variants to disease: scalable approaches to test the functional impact of human mutations. *Hum Mol Genet* 30:R187–R197; PMID 34338757.

Hanna RE (2021) Massively parallel assessment of human variants with base editor screens. *Cell* 184, 1064–1080; PMID 33606977.

Richards S (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17:405–424; PMID 25741868.

Wong AK (2021) Decoding disease: from genomes to networks to phenotypes. *Nat Rev Genet* 22:774–790; PMID 34341555.

Wright CF (2019) Assessing the pathogenicity, penetrance, and expressivity of putative disease-causing variants in a population setting. *Am J Hum Genet* 104(2):275–286; PMID 30665703.

Genetic counseling and prenatal genetic testing

Brezina PR (2012) Preimplantation genetic testing. *Br Med J* 345:e5908; PMID 22990995.

Clarke A (2019) *Harper's Practical Genetic Counselling*, 8th ed. Hodder-Arnold.

Hui L & Bianchi DW (2013) Recent advances in the prenatal interrogation of the human fetal genome. *Trends Genet* 29:84–91; PMID 23158400.

Lo YMD & Chiu RWK (2012) Genomic analysis of fetal nucleic acids in maternal blood. *Annu Rev Genom Hum Genet* 13:285–306; PMID 22657389.

Predictive testing and genetic screening

Umbarger MA (2014) Next-generation carrier screening. *Genet Med* 16:132–140; PMID 23765052.

Wilcken B (2011) Newborn screening: how are we travelling, and where should we be going? *J Inher Metab Dis* 34:569–574; PMID 21499716.

Genomic medicine

Manolio TA (2019) Opportunities, resources, and techniques for implementing genomics in clinical care. *Lancet* 394:511–520; PMID 31395439.

Shendure J (2019) Genomic medicine—progress, pitfalls and promise. *Cell* 177:45–67; PMID 30901547.

Snape K (2019) The new genomic medicine service and implications for patients. *Clin Med* 19(4):273–277; PMID 31308102.

Stark Z (2019) Integrating genomics into healthcare: a global responsibility. *Am J Hum Genet* 104;13–20; PMID 30609404.

Williams MS (2019) Early lessons from the implementation of genomic medicine programs. *Annu Rev Genom Hum Genet* 20:389–411; PMID 30811224.

Wise AL (2019) Genomic medicine for undiagnosed diseases. *Lancet* 394:533–540; PMID 31395441.

Genetic testing of cancers and common genetic disease

Berger MF & Mardis ER (2018) The emerging clinical relevance of genomics in cancer medicine. *Nat Rev Clin Oncol* 15:353–365; PMID 29599476.

Gonzalez de Castro D (2013) Personalized cancer medicine: molecular diagnostics, predictive biomarkers and drug resistance. *Clin Pharmacol Ther* 93:252–259; PMID 23361103.

Mars N (2020) Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med* 26:549–557; PMID 32273609.

Sud A (2021). Will polygenic risk scores for cancer ever be clinically useful? *NPJ Precis Oncol* 5:40; PMID 34021222.

Wald NJ & Old R (2019) The illusion of polygenic disease risk prediction. *Genet Med* 21:1705–1707; PMID 30635622.

Wang L & Wheeler DA (2014) Genome sequencing for cancer diagnosis and therapy. *Annu Rev Med* 65, 33–48; PMID 24274147.

Ethical issues in genetic testing

Caulfield T & McGuire AL (2012) Direct-to-consumer genetic testing: perceptions, problems and policy responses. *Annu Rev Med* 63:23–33; PMID 21888511.

de Jong A (2011) Advances in prenatal screening: the ethical dimension. *Nat Rev Genet* 12:657–663; PMID 21850045.

Dheensa S (2017) Approaching confidentiality at a familial level in genomic medicine: a focus group study with healthcare professionals. *BMJ Open* 7:e012443; PMID 28159847.

Lucassen A & Gilbar R (2018) Alerting relatives about heritable risks: the limits of confidentiality. *BMJ* 361:k1409; PMID 29622529.

Hollands GJ (2016) The impact of communicating genetic risks of disease on risk reducing health behaviour: systematic review with meta-analysis. *BMJ* 352:i1102; PMID 26979548.

Manrai AK (2016) Genetic misdiagnoses and the potential for health disparities. *N Engl J Med* 375:655–665; PMID 27532831.

Marcon AR (2018) Representing a "revolution": how the popular press has portrayed personalized medicine. *Genet Med* 20:950–956; PMID 29300377.

Ross LF (2013) Technical report: ethical and policy issues in genetic testing and screening of children. *Genet Med* 15:234–245; PMID 23429433.

Clinical and public health genomics: challenges and ethics

Ackerman JP (2016) The promise and peril of precision medicine: phenotyping still matters most. *Mayo Clin Proc* 91:1606–1616; PMID 27810088.

Bertier G (2018) Is it research or is it clinical? Revisiting an old frontier through the lens of next-generation sequencing technologies. *Eur J Med Genet* 61:634–641; PMID 29704685.

Dheensa S (2018) Towards a national genomics medicine service: the challenges facing clinical-research hybrid practices and the case of the 100 000 genomes project. *J Med Ethics* 44(6):397–403.

Faden RR (2013) An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. *Hastings Cent Rep* 43:S16–S27; PMID 23315888.

Jackson L (2021) *Use of SNP chips to detect rare pathogenic variants: retrospective, population based diagnostic evaluation BMJ* 372:n214; PMID 33589468.

Johnson SB (2020). Rethinking the ethical principles of genomic medicine services. *Eur J Hum Genet* 28:147–154; PMID 31534213.

McEwen JE (2013) Evolving approaches to the ethical management of genome data. *Trends Genet* 29:375–382; PMID 23453621.

Miller DT (2021) *Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2021 update: a policy statement of the American College of Medical Genetics and Genomics* Genet Med 23(8):1391–1398; PMID 34012069.

Milne R (2021) Demonstrating trustworthiness when collecting and sharing genomic data: public views across 22 countries. *Genome Med* 13:92; PMID 34034801.

Nuffield Council on Bioethics (2018) *Genome Editing and Human Reproduction: Social and Ethical Issues*. Nuffield Council on Bioethics, London.

Popejoy AB (2018) The clinical imperative for inclusivity: race, ethnicity, and ancestry (REA) in genomics. *Hum Mutat* 39:1713–1720; PMID 30311373.

Wright CF (2019) Genomic variant sharing: a position statement. *Wellcome Open Res* 4:22; PMID 31886409.

Wright CF (2019) When genomic medicine reveals misattributed genetic relationships-the debate about disclosure revisited. *Genet Med* 21:97–101; PMID 29904162.

Glossary

3′end
The end of a DNA or RNA strand that is linked to the rest of the chain only by carbon 5′ of the sugar, not carbon 3′ (Box 1.1, Figure 1).

5′end
The end of a DNA or RNA strand that is linked to the rest of the chain only by carbon 3′ of the sugar, not carbon 5′ (Box 1.1, Figure 1).

adaptive immunity/adaptive immune system
Specific immune responses that rely on the recognition of foreign antigen by antibodies and T-cell receptors.

allele frequency
The frequency of an allele in a population; that is, the proportion of all alleles at a locus that are the allele in question (often inaccurately represented as gene frequency).

allele
Individual version of a gene or DNA sequence at a locus on a single chromosome; often also used loosely to describe genetic variants at the protein level.

allogeneic
Describing cell and organ transplantation (or the transplanted cells) in which the donor cells are genetically different from that of the recipient. Compare *autologous*.

amino acid
The fundamental repeating unit of a polypeptide; a building block for a protein (Figure 2.2 and Table 7.3).

amplification
1. An artificial increase in DNA sequence copy number as a result of *cloning* or *PCR* (Section 3.1). 2. A natural increase in gene copy number in response to natural selection in organisms (Figure 4.8) or tumors (Figure 10.7).

anaphase lag
Loss of a chromosome because it moves too slowly at anaphase to get incorporated into a daughter nucleus.

aneuploidy
A chromosome constitution with one or more chromosomes extra or missing from a full (euploid) set – see pp. 211–2.

angiogenesis
Process whereby new blood vessels are formed by sprouting from existing vessels.

annealing
Process whereby two single-stranded nucleic acids form a stable double-stranded nucleic acid by *base pairing*. The reverse of denaturation.

anticipation
The tendency for the severity of a condition to increase in successive generations (p. 129). Commonly due to bias of ascertainment, but a genuine outcome in the case of some *dynamic mutations*.

antigen
A molecule that can induce an adaptive immune response or that can bind to an antibody or T-cell receptor.

antigen presentation

The process by which antigen is presented in combination with an MHC (HLA) protein on the surface of certain cells so that it can be recognized by receptors on lymphocytes (Section 4.4 and Box 8.3Figure 1).

antisense RNA

An RNA transcript that has a *complementary sequence* to a mRNA (or some functional noncoding RNA). Naturally occurring antisense RNAs, made using the non-template strand of a gene, are important regulators of gene expression.

antisense (or template) strand

The DNA strand of a gene that, during transcription, is used as a template by RNA polymerase for the synthesis of mRNA (Figure 2.1).

apoptosis

A natural way of getting rid of unwanted or diseased cells in which the cell is targeted for destruction by various stimuli. Rapid fragmentation of the cell follows, after which the resulting cell fragments are phagocytosed by neighboring cells.

association

A tendency of two *characters* (such as diseases or marker alleles) to occur together at nonrandom frequencies. Association is a simple statistical observation, not a genetic phenomenon, but can be caused by *linkage disequilibrium* (Section 8.2).

autoimmune disorders

Diseases that arise because the distinction between self and nonself fails so that the body mounts an abnormal immune response against one or more self molecules.

autologous

Describing cells or tissues that were obtained from or pertain to the same individual.

autosome
Any chromosome other than the sex chromosomes, X and Y.

autozygosity
In an inbred person, homozygosity for alleles identical by descent.

balancing selection
Selection working simultaneously in opposite directions on the same variant; can result in heterozygotes for a harmful mutation having a higher biological *fitness* than normal homozygotes (p. 136).

base complementarity
The relationship between bases on opposite strands of a double-stranded nucleic acid: A always occurs opposite T (or U in RNA) and G always occurs opposite C in DNA (but in RNA, G sometimes base pairs with U).

base pair/base pairing
The outcome/process of stable hydrogen bonding between two complementary bases, a purine and a pyrimidine. The bases may reside on opposing strands of a duplex nucleic acid (Figure 1.4), or on the same RNA strand (Figure 2.4A). Transient DNA-RNA and RNA-RNA base pairing can allow functional interaction between different molecules.

benign tumor
An abnormal cell growth that is confined to a specific site within a tissue and shows no evidence of invading adjacent tissue.

biologic
Any biological drug, such as a therapeutic monoclonal antibody or recombinant protein.

biomarker

Any characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.

biotin–streptavidin system

A tool for isolating labeled molecules. The bacterial protein streptavidin happens to bind biotin (vitamin B7) with exceptionally high affinity. Biotinylated molecules can be isolated by using streptavidin-coated magnetic beads (Figure 8.7).

blastocyst

An embryo at a very early stage of development when it consists of a hollow ball of cells with a fluid-filled internal compartment (Figure 2 in Box 9.2).

blastomere

One of the multiple cells formed when the fertilized egg undergoes cleavage divisions.

boundary elements

Regulatory DNA sequences that define the boundary between differentially regulated loci by limiting or opposing the action of enhancer elements

capping

A stage in RNA processing. A special nucleotide, 7-methylguanosine triphosphate, is joined by a 5′–5′ phosphodiester bond to the 5′ end of a primary transcript. Capping is important for the stability of the RNA.

cancer

1. One of a heterogeneous group of disorders whose common features are uncontrolled cell growth and cell spreading. 2. A tumor that has become *malignant*.

carrier
A person, usually asymptomatic, who carries a genetic variant that can cause disease after being transmitted to the next generation, or that can contribute to disease in later life.

case-control study
A study in which samples from affected individuals (cases) are analyzed and compared with equivalent samples from unaffected control individuals.

cDNA (complementary DNA)
DNA synthesized by the enzyme reverse transcriptase using RNA (often mRNA) as a template.

centromere
The primary constriction of a chromosome, separating the short arm from the long arm, and the point at which spindle fibers attach to pull chromatids apart during cell division.

character (or trait)
An observable property of an individual, such as eye color or ABO blood group type.

chimera
An organism derived from more than one zygote.

chromatid
One of a pair of sister chromatids that form when a chromosome replicates and persist until the anaphase stage of mitosis (see Figure 1.10).

chromatin
The nucleoprotein material of a chromosome.

chromatin remodeling

Movement, dissociation, or reconstitution of nucleosomes in chromatin, as part of the systems controlling chromatin conformation.

chromosomal microarray analysis
Clinical application of *microarray hybridization*. The usual object is to scan a genomic DNA sample for changes in copy number (deletions or duplications) of large DNA segments.

chromosome
In eukaryotes, a nucleoprotein structure formed when a nuclear DNA molecule is complexed with various types of proteins and occasionally some RNAs. The complexing helps compact the immensely long DNA molecules.

***cis*-acting**
(of gene regulation by short sequence elements) Term used to describe any gene regulation in which a regulatory DNA or RNA sequence controls the expression of some other sequence present on the *same* nucleic acid molecule (Figures 6.1, 6.2).

***cis*-acting RNA**
a type of regulatory long noncoding antisense RNA that remains attached to the DNA strand from which it is transcribed but can base pair with the sense strand transcribed from the opposing DNA strand of the same DNA molecule, and thereby regulate its expression.

clones/cloning
Identical copies (of a DNA sequence, a cell, or an organism)/process of making the same. In genetic research, this often means cells containing identical recombinant DNA molecules.

CNV
See *copy number variation*.

coding DNA
A segment of DNA whose sequence is used directly to specify a polypeptide (via a mRNA).

co-dominant
Term used to describe a heterozygous state in which both alleles are fully expressed.

codon
A sequence of three nucleotides (strictly in mRNA, but by extension, in genomic coding DNA) that specifies an amino acid or a translation stop signal.

coefficient of inbreeding
The proportion of loci at which a person is homozygous by virtue of the consanguinity of their parents (Section 5.2).

coefficient of relationship
Of two people, the proportion of loci at which they share alleles identical by descent (Box 5.2).

complementary sequences (or strands)
Nucleic acid sequences (or strands) that can form a stable double-stranded nucleic acid by *base pairing*.

complementary DNA
See *cDNA*.

compound heterozygote
A person with two different mutant alleles at a locus.

conformation
Of a complex molecule, the three-dimensional shape—the result of the combined effects of many weak noncovalent bonds.

consanguineous

Description of persons who are closely related because they have descended from a very recent common ancestor (often within the previous three or four generations), usually as a result of a marriage between cousins.

conservative substitution

A nucleotide substitution that changes a codon so that it makes a different, but chemically similar, amino acid.

conserved sequence

DNA or amino acid sequence that is identical or recognizably similar across a range of organisms, suggestive of an important function.

constitutional

(of genetic variation, mutation, chromosome abnormality) Present in the genetic material of the zygote, and therefore present in every nucleated cell of a person.

constitutive heterochromatin

Heterochromatin that remains condensed throughout the cell cycle. Found at centromeres plus some other regions. See Box 2.3, Figure 2.8.

copy number variation (CNV)

Variation between individuals in the number of copies in their genomes of a specific, moderately long to large DNA sequence (from hundreds of base pairs to many megabases). The term CNV is also used to denote a rare copy number variant (frequency less than 1%); if the frequency is above 1%, copy number polymorphism (CNP) is often used (pp. 91–2).

CpG island

Short stretch of DNA, often less than 1 kb long, containing frequent unmethylated CpG dinucleotides. CpG islands tend to mark the 5′ ends of genes (Box 6.1).

CRISPR-Cas

A type of natural prokaryotic adaptive immunity. Adapted as a *genome editing* technique that uses artificial RNA *guide sequences*. See Figures 9.22, 9.24.

cryptic splice site

A sequence in pre-mRNA with significant homology to a splice site. Cryptic splice sites may be used as splice sites when splicing is disturbed or after a base substitution mutation that increases the resemblance to a normal splice site (Figure 7.4).

cross-linking

(in DNA) Abnormal occurrence of covalent bonds directly linking two bases. The cross-linked bases may be on the same strand or on opposite strands (Figure 4.1). In proteins, the disulfide bond is a natural form of cross-linking (Figure 2.5).

crossover

An act of meiotic *recombination*, or the physical manifestation of that (as seen under the microscope) (Figures 1.13 and 1.14).

cytokines

Extracellular signaling proteins or peptides that act as local mediators in cell–cell communication.

dedifferentiation

Epigenetic reprogramming of a *differentiated cell* so that the cell becomes less specialized (Box 9.1).

denaturation

1. Dissociation of double-stranded nucleic acid to give single strands. 2. Destruction of the three-dimensional structure of a protein by heat or high pH.

derivative chromosome
A chromosome that has been structurally rearranged, for example by translocation, but retains a centromere.

differentiation
(of a cell) Natural process of epigenetic modification that causes a cell to become more specialized.

diploid
Having two copies of each type of chromosome; the normal constitution of most human somatic cells.

direct repeats
Two or more copies of a sequence that occur in the same 5′ → 3′ direction on a single DNA strand. Usually used to mean repeats that are separated on the DNA; repeats that are directly adjacent to one another are normally described as *tandem repeats*.

distal
(of chromosome) Comparatively distant from the centromere (Box 7.2).

DNA libraries
The result of cloning random DNA fragments or molecules to produce a collection of cells containing different recombinant DNAs (which must then be screened to find any desired sequence).

dominant
In human genetics, any trait that is expressed in a heterozygote.

dominant-negative effect
The situation in which a mutant protein antagonizes the function of its normal counterpart in a heterozygous person (Figure 7.19).

dosage-sensitive gene

A chromosomal gene that, when present in one copy instead of the normal two copies is associated with disease (*haploinsufficiency*), or that can cause disease when overexpressed (Box 7.3).

driver mutations

In cancer, mutations that assist tumor development, being subject to positive selection during tumorigenesis as opposed to passenger mutations in tumorigenesis (which are not positively selected or causally implicated in cancer development.

duplex

A double-stranded nucleic acid.

dynamic mutation

An unstable expanded repeat that changes in size between parent and child (Section 7.3).

embryonic stem (ES) cell line

Embryonic stem cells that have continued to proliferate after subculturing for a period of 6 months or longer and that are judged to be *pluripotent* and genetically normal.

endonuclease

An enzyme that cuts DNA or RNA at an internal position in the chain.

enhancer

A set of clustered short sequence elements that stimulate the transcription of a gene and whose function is not critically dependent on their precise position or orientation (Section 6.1).

epigenetic

Heritable (from mother cell to daughter cell, or sometimes from parent to offspring), but not produced by a change in DNA sequence.

epigenetic marks (or settings)

Patterns of epigenetic modification, such as DNA methylation, histone modification, and nucleosome spacing patterns that permit chromatin to switch between open (transcriptionally active) and condensed (transcriptionally inactive) forms.

epigenome

The totality of epigenetic marks in a cell.

epimutation

A change in chromatin organization causing a change in expression of one or more genes *without any change to the DNA sequence* (Figure 6.21). Can be induced by mutation at a distant gene locus regulating chromatin modification, by environmental factors (resulting in metabolic changes or inflammation, for example), and certain chromosome abnormalities (*position effects*).

episome

Any DNA sequence that can exist in an autonomous (self-replicating) extrachromosomal form in the cell.

epistasis

Literally ‘standing above’. Gene A is epistatic to gene B if A functions upstream of B in a common pathway. Loss of function of A will cause all the effects of loss of function of B, and maybe other effects as well.

epitope

The part of an immunogenic molecule to which an antibody responds.

euchromatin

The fraction of the nuclear genome that contains transcriptionally active DNA and that, unlike heterochromatin, adopts a relatively extended conformation.

exon

Originally, any segment of an RNA transcript that is retained during RNA *splicing*, but now used widely to mean the corresponding sequence in genomic DNA. Individual exons may contain coding sequences that are translated and/or noncoding sequences (Figure 2.1).

exome

The totality of exons in a genome.

exon shuffling

An evolutionary process in which exons from one gene are copied and inserted into a different gene (Figure 2.16).

exon skipping

Occasional failure to include an exon within an RNA transcript (Figure 6.7).

exonuclease

An enzyme that digests a DNA or RNA strand from one end. It may be a 3′ or 5′ exonuclease.

facultative heterochromatin

Heterochromatin that may reversibly decondense to form euchromatin, depending on the requirements of the cell. See Box 2.3.

FISH

See *fluorescence in situ hybridization*.

fitness (*f*)

In population genetics, a measure of the success in transmitting genotypes to the next generation, relative to the most successful genotype. Also called biological or reproductive fitness. *f* always lies between 0 and 1.

fluorescence *in situ* hybridization (FISH)

Hybridization of a fluorescently labeled probe to the denatured DNA of chromosome preparations that have been immobilized on a solid surface (Figures 11.4), or to the RNA of similarly immobilized cells.

fluorophore (or fluorochrome)
A fluorescent chemical group, used for labeling nucleic acids or proteins (Box 3.2).

founder effect
High frequency of a particular allele in a population because the population is derived from a small number of founders, one or more of whom carried that allele.

fragile site
Location on a chromosome where the chromatin of metaphase chromosomes appears condensed under certain culture conditions. Most examples do not cause disease.

frameshift
A change in the base sequence of coding DNA that removes or adds nucleotides so as to change the translational reading frame (Box 2.1).

G-banding
The standard method of identifying chromosomes under the microscope. *See* Figure 2.8 for an example.

gain-of-function mutations
Mutations that cause the gene product to do something abnormal, rather than simply to lose function. Usually the gain is a change in the timing or level of expression (Sections 7.7 and 10.2).

gamete
Sperm or egg; a haploid cell formed when a germ cell precursor undergoes meiosis.

gene

1. A functional DNA that is used to make a valuable RNA or protein end-product. 2. A factor that controls a phenotype and segregates in pedigrees according to Mendel's laws.

gene conversion

A naturally occurring nonreciprocal genetic exchange in which a short sequence of one DNA strand is altered so as to become identical to the sequence of another DNA strand (Figure 7.9).

gene dosage

The copy number of a gene. Alteration of the normal gene copy number causes reduced expression (too little gene product) or overexpression (too much product). For *dosage-sensitive genes* (Box 7.3), the amount of gene product made is critically important.

gene editing/genome editing

Making desired changes to a *specific* gene (or other target) sequence in the genome of *intact* cultured cells. *See* also *gene targeting* and *CRISPR-Cas*.

gene family

A set of related genes that arose by some type of gene duplication (Section 2.5).

gene frequency

see *allele frequency*.

gene knockout

The targeted inactivation of a predetermined gene within intact cells so as to artificially create a *null allele*.

gene pool

All the genes (in the whole genome or at a specified locus) in a particular population.

gene silencing/gene suppression

Gross or significant reduction in gene expression occurring naturally by altered epigenetic settings, and that can occur both naturally and artificially through *RNA interference*.

gene targeting

A type of gene editing using homologous recombination to specifically alter a pre-determined gene of interest within intact cells (Box 9.2, Figure 2).

gene therapy

Treating disease by genetically modifying the cells of a patient. May involve adding a functional copy of a gene that has lost its function, inhibiting a gene showing a pathological gain of function, or, more generally, replacing a defective gene.

genetic background

The genotypes at all loci other than one locus under active investigation. Variations in genetic background (*modifier genes*) are a major reason for imperfect genotype–phenotype correlations (Section 7.9).

genetic code

The relationship between a codon and the amino acid it specifies (Figure 7.2).

genetic counseling

The process in which one or more members of a family that have, or are at risk of developing or transmitting, an inherited disease are informed by health professionals of the consequences and nature of the disorder, the probability of developing or transmitting it, and the options open to them.

genetic drift

Random changes in allele frequencies over generations because of random fluctuations in the proportions of the alleles in the parental population that are transmitted to offspring. Only significant in small populations.

genetic redundancy

Partly or completely overlapping function of genes at more than one locus, so that *loss-of-function mutations* at one locus do not cause overall loss of function.

genome/genomics

The total set of different DNA molecules of a cell, organelle, or organism/study of the same. The human genome consists of 24 different chromosomal DNA molecules and one mitochondrial DNA molecule.

genomic constraint

The constraint that *natural selection* imposes on variation at functionally important DNA sequences

genome browser

A computer program that provides a graphical interface for interrogating genome databases.

genome editing

Artificial manipulation of an intact cell that is designed to make a double-strand break at just one locus and subsequently to make a desired change to the base sequence at that locus. See Figure 9.24 for an example.

genome (or gene) imprinting

See *imprinting*.

genomewide association (study)or GWA(S)

The standard approach to identifying factors governing susceptibility to complex disease (Figure 8.15).

genotype
The genetic constitution of an individual, either overall or at a specific locus.

germ line
The germ cells (gametes) and those cells that give rise to them; other cells of the body constitute the soma.

germ-line (or gonadal) mosaic
An individual who has a subset of germ-line cells carrying a mutation that is not found in other germ-line cells.

guide RNA
A short RNA that can base pair with a specific target sequence in order to guide some DNA-targeting (or RNA-targeting) enzyme to recognize the target sequence.

haploid
Term used to describe a cell (typically a gamete) that has only a single copy of each chromosome (for example the 23 chromosomes in a human sperm or egg).

haploinsufficiency
A locus shows haploinsufficiency if producing a normal phenotype requires more gene product than the amount produced by a single functional allele (Box 7.3).

haplotype
A series of alleles found at linked loci on a single chromosome (Box 4.3 and Figure 8.2).

haplotype block

A region of DNA showing limited haplotype diversity (Box 8.4).

Hardy–Weinberg law (or equilibrium)
The simple relationship between allele frequencies and genotype frequencies that is found in a population under ideal conditions (Section 5.4).

hemizygous
Having only one copy of a gene or DNA sequence in diploid cells. Males are hemizygous for most genes on the sex chromosomes. Deletions occurring on one autosome produce hemizygosity in males and in females.

heritability
The proportion of the causation of a character that is due to genetic causes (Section 8.2).

heterochromatin
Highly condensed chromatin showing little or no evidence of active gene expression. Facultative heterochromatin may reversibly decondense to form *euchromatin*, depending on the requirements of the cell, but *constitutive heterochromatin* remains condensed throughout the cell cycle.

heteroduplex
Double-stranded DNA in which there is some mismatch between the two strands.

heteroplasmy
Mosaicism, usually within a single cell, for mitochondrial DNA variants (Section 7.6).

heterozygous/heterozygote

Having two different alleles at a particular locus/an individual with this property.

heterozygote advantage

The situation when a person heterozygous for a mutation has a reproductive advantage over both homozygotes for this mutation and also normal homozygotes. Sometimes called overdominance. Heterozygote advantage is one reason why severe recessive diseases may remain common (Section 5.4).

homologs (homologous chromosomes)

The two copies of a chromosome in a diploid cell. Unlike sister chromatids, homologous chromosomes are not copies of each other: one was inherited from the father and the other from the mother.

homologs (genes)

Two or more genes whose sequences are significantly related because of a close evolutionary relationship. They include *orthologs*, equivalent genes in two or more species that evolved from a single gene present in a common evolutionary ancestor, and paralogs that evolved by gene duplication such as the two α-globin genes present in humans.

homoplasmy

Of a cell or organism, having all copies of the mitochondrial DNA identical, as opposed to *heteroplasmy*.

homozygous/homozygote

Having two identical alleles at a particular locus/a person with this property. For clinical purposes a person is often described as homozygous *AA* if they have two normally functioning alleles, or homozygous *aa* if they have two pathogenic alleles at a locus, regardless of whether the alleles are in fact completely identical at the DNA sequence level. See also *autozygosity*.

hybridization
(of nucleic acids and oligonucleotides) Process in which complementary single strands are allowed to base pair (*anneal*) to form duplexes.

hybridization stringency
The degree to which the conditions (temperature, salt concentration, and so on) during a hybridization assay permit sequences with some mismatches to hybridize. High stringency conditions allow perfect matches only (Figure 3.7).

immunotherapy
Traditionally, type of therapy that uses substances to stimulate or suppress the immune system so as to help the body to fight cancer or other diseases, but now including the use of genetically engineered antibodies and T cells.

imprinting
(of certain mammalian genes) An epigenetic phenomenon in which the expression of the gene is determined by its parental origin (pp.160–3).

indels
Insertion/deletion variants, often involving a single nucleotide, but sometimes involving more nucleotides. (The definition is a little imprecise; in practice it usually includes variants that differ by possessing/lacking a sequence of up to 50 nucleotides.)

induced pluripotent stem (iPS) cells
Somatic cells that have been treated with specific genes, gene products, or other agents to reprogram them to resemble pluripotent stem cells. They can then be induced to differentiate into desired cell types (Box 9.1).

innate immunity/innate immune system

System of nonspecific response to a pathogen using the natural defenses of the body, as opposed to the *adaptive immunity/adaptive immune system*.

inner cell mass (ICM)

A group of cells located internally within the blastocyst which will give rise to the embryo proper (Box 9.1, Figure 2).

insulator

DNA element that act as a barrier to the spread of chromatin changes or the influence of *cis*-acting elements.

interphase

All the time in the cell cycle when a cell is not dividing.

intron

Originally any segment of a transcript that is cut out and discarded during RNA *splicing*, but now widely used to mean the corresponding sequence in genomic DNA (Figure 2.1).

isochromosome

An abnormal symmetrical chromosome consisting of two identical arms, either the short arm or the long arm of a normal chromosome.

isoform

Alternative form of a protein as a result of differential expression of the same gene or through the production of different but highly related proteins from two or more loci.

karyotype

A summary of the chromosome constitution of a cell or person, such as 46,XY, but widely used loosely to mean an image showing the chromosomes of a cell sorted in order and arranged in pairs.

ligand

Any molecule that binds specifically to a receptor or other molecule, such as the trimeric FASLG ligand that binds to the FAS receptor in Figure 10.13..

ligase

DNA ligase is an enzyme that can seal single-strand *nicks* in double-stranded DNA or covalently join two oligonucleotides that are hybridized at adjacent positions on a DNA strand.

lineage

(of cells) In development, the ancestry and descendants of a cell, as traced backward or forward through successive cell divisions.

linkage analysis

Any statistical method that aims to identify chromosomal regions that co-segregate with a disease gene, or other gene of interest.

linkage disequilibrium

A statistical association between particular alleles at separate but linked loci, normally the result of a particular ancestral haplotype being common in the population studied. An important tool for high-resolution mapping (Section 8.2).

liposome

A synthetic lipid vesicle that can be used to introduce DNA into cells.

liquid biopsy

A test done on a blood sample to look for cancer cells from a tumor that are circulating in the blood, or for pieces of DNA from tumor cells in the blood.

locus (plural: loci)

A unique chromosomal location defining the position of an individual gene or DNA sequence.

lod score (*Z*)
A measure of the likelihood of genetic linkage between loci. The log (base 10) of the odds that the loci are linked (with recombination fraction *q*) rather than unlinked. For Mendelian characters a lod score greater than +3 provides minimal evidence of linkage; one that is less than –2 is evidence against linkage (Box 8.1).

loss-of-function mutations
Mutations that cause a gene product to lose its function, partly or totally (Section 7.7).

loss of heterozygosity (LOH)
Homozygosity or hemizygosity in a tumor or other somatic cell when the constitutional genotype is heterozygous. Evidence of a somatic genetic change (Section 10.2 and Figure 10.11).

major histocompatibility complex (MHC)
A large gene cluster containing multiple genes including, notably, genes that function in antigen recognition by binding fragments of antigens and presenting them on the surface of T cells. The human MHC is known as the HLA complex (see Boxes 4.4 and 8.3). *See also* MHC restriction.

malignant tumor
A tumor whose cells show evidence of spreading (invading adjacent tissue, disseminating through the bloodstream and/or lymphatic system).

marker
(molecular) A chemical group or molecule that can be assayed in some way.

meiosis
The specialized reductive form of cell division used exclusively to produce gametes (Figures 1.13 and 1.14).

Mendelian
Description for a character whose pattern of inheritance suggests it is caused by variation at a single chromosomal locus.

mesenchyme
Connective tissues.

messenger RNA (mRNA)
A processed gene transcript that carries protein-coding information to cytoplasmic ribosomes.

meta-analysis
A statistical analysis of combined data from a number of independent studies of the same topic.

metastasis
The process whereby cells from a primary malignant tumor are disseminated via the blood stream or lymphatic system to establish secondary tumors at distant sites in the body.

MHC restriction
The requirement that when a T cell is confronted with a complex of a self-MHC molecule and a foreign peptide antigen bound to it, it will only respond to the antigen when it is bound to a *particular* MHC molecule (Box 8.3, Figure 1).

microarray hybridization
A nucleic acid hybridization assay in which thousands to millions of different oligonucleotide (or DNA) probes are fixed at specific grid coordinates on a miniature solid surface and allowed to hybridize to complementary sequences within a solution containing a heterogeneous test sample population of labeled DNA or RNA molecules (Figure 3.9).

microbiome (or microbiota)

The aggregate of microorganisms that share our body space; most of them are found in the gastrointestinal tract.

microRNAs (miRNAs)
Short (21–22-nucleotide) RNA molecules encoded within normal genomes that have a major role in the regulation of gene expression (Figure 6.10).

microsatellite
Small array of *tandemrepeats* of a very simple DNA sequence, usually 1–4 bp, for example $(CA)_n$. The total length of the array is usually less than 0.1 kb. A polymorphic microsatellite is alternatively known as a short tandem repeat polymorphism (Figure 4.6).

mismatch repair
A form of DNA repair in which very simple DNA replication errors (nucleotide substitutions and deletions/insertions of one or two nucleotides) are repaired (Figure 10.16).

missense mutations
Changes in a coding sequence that cause one amino acid in the gene product to be replaced by a different one (Section 7.2).

mitosis
The normal process of cell division, which usually produces daughter cells genetically identical to the parent cell (Figure 1.12).

modifier (gene)
A gene whose expression can influence a phenotype resulting from a mutation at another locus (Section 7.9).

monozygotic
Originating from a single zygote, as in identical twins (other twins are dizygotic, having originated from different zygotes).

mosaic
An individual who has two or more genetically different cell lines derived from a single zygote. The difference may be point mutations, large-scale mutations or chromosomal abnormalities (Box 5.3).

mRNA
See *messenger RNA*.

mtDNA
Mitochondrial DNA (Figure 2.12).

multifactorial
A character that is determined by some unspecified combination of genetic and environmental factors.

mutagen
An agent that results in an increased mutation frequency.

mutation
1. A localized change in the base sequence of a DNA molecule. 2. The process that creates it.

mutation scanning/screening
Testing for any *undefined* change in the base sequence of a genome or genome component (notably exon, gene, or exome) in the hope of identifying abnormal variants correlating with disease. As opposed to testing for a *specific* DNA variant.

natural selection
Process whereby the population frequencies of alleles change by causing a change in the biological *fitness* of the individuals who carry them. Many alleles cause reduced biological fitness (*purifying* or *negative selection*); a few alleles cause increased biological fitness of the

individuals who carry them (*positive selection*). See also *balancing selection*.

ncRNA
See *noncoding RNA*.

next generation sequencing
(also called *massively parallel sequencing*) Any method that permits very high-throughput DNA sequencing by sequencing many molecules in parallel. See Box 11.1.

nick (in DNA)
Cleavage of a single phosphodiester bond on one DNA strand only.

non-allelic homologous recombination (NAHR)
Recombination between misaligned DNA repeats, either on the same chromosome, on sister chromatids or on homologous chromosomes. NAHR generates recurrent deletions, duplications, or inversions (Section 7.4).

noncoding RNA (ncRNA)
mature RNA transcript that is not translated to make a polypeptide (Figure 2.7).

nondisjunction
Failure of chromosomes (sister chromatids in mitosis or meiosis II; paired homologs in meiosis I) to separate (disjoin) at anaphase (Figure 7.16). The major cause of numerical chromosome abnormalities.

nonhomologous end joining
Form of repair of double-strand breaks in DNA that involves the fusion of broken ends without copying from a DNA template.

non-penetrance

The situation when somebody carrying an allele that normally causes a phenotype to be expressed does not show that phenotype, as a result of interaction with alleles of other genes (*modifier genes*) or with non-genetic factors (Figure 5.12).

nonsense mutation

A nucleotide substitution that changes a codon specifying an amino acid so that it becomes a premature termination codon (Section 7.2).

nonsense-mediated mRNA decay

A cellular mechanism that degrades mRNA molecules containing a premature termination codon more than 50 nucleotides upstream of the last splice junction (Box 7.1). A stop codon less than 50 nucleotides from the last splice junction may often be harmless, but sometimes a short toxic polypeptide may be produced.

nonsynonymous substitution (or mutation)

A change in the sequence of a codon that results in a different codon interpretation. Table 7.2 gives the different classes.

nucleosome

The basic structural unit of chromatin, comprising 146 bp of DNA wound around an octamer of histone molecules (Figures 1.7 and 6.13).

nucleotide

The fundamental repeating unit of a nucleic acid, consisting of a sugar to which is covalently attached a base and a phosphate group (Figure 1.2).

null allele

Any mutant allele where the normal gene product is not made or is completely non-functional.

odds ratio

In case-control studies, the relative odds of a person with or without a factor under study being a case (Table 8.6).

OMIM

Online Mendelian Inheritance in Man database at https://www.ncbi.nlm.nih.gov/omim

oncogene

A gene that when activated in some way (often by a change that stimulates its expression) can help to transform a normal cell into a tumor cell. Originally the word was reserved for activated forms of the gene (while the normal unactivated cellular gene was called a proto-oncogene), but this distinction is now widely ignored.

open reading frame

A continuous sequence of *coding DNA*.

origin of replication

A site on a DNA molecule where replication can be initiated.

orthologs

Homologous genes present in different organisms having descended from a common ancestral gene.

PCR (polymerase chain reaction)

The standard technique used to amplify short DNA sequences (Figure 3.3).

penetrance

The frequency with which a genotype manifests itself in a given phenotype.

pedigree

A limited family tree; a more extensive family tree is a kindred.

personalized medicine
A model of health care in which medical decisions and practice are tailored to the individual patient. Knowledge of a person's genome, for example, can allow more informed decisions about the suitability of prescribing certain drugs, and knowledge of cancer mutations may allow suitably targeted therapies.

pharmacodynamics
The study of the response of a target organ or cell to a drug.

pharmacogenetics
The study of the influence of individual genes or alleles on the metabolism or function of drugs.

pharmacokinetics
The study of the absorption, activation, catabolism, and elimination of a drug.

phasing
Converting genotypes into haplotypes in genome wide association studies..

phenocopy
A person or organism that has a phenotype normally caused by a certain genotype but does not have that genotype. Phenocopies may be the result of a different genetic variant, or of an environmental factor.

phenome
The totality of phenotypes of an individual organism.

phenotype
The observable characteristics of a cell or organism, including the result of any test that is not a direct test of the genotype.

phosphodiester bond

The link between adjacent nucleotides in DNA or RNA.

plasmid
A small circular DNA molecule that can replicate independently in a cell. Modified plasmids are widely used as cloning vectors (Section 3.1).

pleiotropy
The common situation in which variation in one gene affects several different aspects of the phenotype.

ploidy
The number of complete sets of chromosomes in a cell. Gametes are *haploid* and most normal cells are *diploid*, but some of our cells naturally have multiple chromosome sets (polyploidy) or none at all (nulliploidy).

pluripotent (of a mammalian stem cell)
Capable of giving rise to descendant cells that participate in the formation of all of the tissues of an embryo except the extraembryonic membranes.

PMID
PubMed identifier, a seven-digit or eight-digit number that, when typed into the query box at the NCBI PubMed database (http://www.ncbi.nlm.nih.gov/pubmed/), allows electronic access to a specific article in a biomedical journal.

point mutation
A mutation causing a small alteration in the DNA sequence at a locus, often changing just a single nucleotide.

polyadenylation/poly(A) tail
Addition of 200 or so adenosines to the 3′ end of a mRNA. The resulting poly(A) tail is important for stabilizing mRNA (Section 2.1).

polygenic

Description of a character determined by the combined action of a number of genetic loci. Polygenic theory (Box 8.2) assumes that there are very many loci, each with a small effect.

polygenic risk score

Assessment of the risk of a specific condition based on the collective influence of very many genetic variants including variants not known to be associated with genes relevant to the condition.

polymorphism

The existence of two or more variants (alleles, phenotypes, sequence variants, chromosome structure variants) at significant frequencies in the population. Often also used more loosely to mean any sequence variant present at a frequency of more than 1% in a population.

polypeptide

A string of amino acids linked by peptide bonds. Proteins may contain one or more polypeptide chains.

positional cloning

Identifying a disease gene using knowledge of its chromosomal location.

position effect

Complete or partial silencing of a gene after some chromosome rearrangement that results in the gene becoming heterochromatinized – see Figure 6.22.

positive selection

Selection in favor of a particular genotype that confers increased biological *fitness* (Section 4.3).

potency

Of a cell, its potential for dividing into different cell types. Cells can be totipotent, *pluripotent*, multipotent, or committed to one fate.

premutation allele
Among diseases caused by *dynamic mutations*, a repeat expansion that is large enough to be unstable on transmission but not large enough to cause disease.

primary structure
Of a polypeptide or nucleic acid, the linear sequence of amino acids or nucleotides in the molecule.

primary transcript
The RNA product of transcription of a gene by RNA polymerase, before splicing. The primary transcript of a gene contains all the exons and introns.

primer
A short oligonucleotide, often 16–25 bases long, which base pairs specifically to a target sequence to allow a polymerase to initiate the synthesis of a complementary strand.

primordial germ cells
Cells in the embryo and fetus that will ultimately give rise to germ-line cells.

probe
Known DNA or RNA fragments used in a hybridization assay to identify closely related target sequences within a complex, poorly understood population of nucleic acid molecules (the test sample)—see Section 3.3.

prodrug
An inactive precursor to a therapeutic drug that is administered to a patient and activated within the body after natural conversion by a drug-metabolizing enzyme or other component (Section 9.2).

promoter

A combination of short sequence elements, usually just upstream of a gene, to which RNA polymerase binds so as to initiate transcription of the gene (Figure 6.1).

protective factor

A variant that reduces susceptibility to disease (Table 8.11).

proofreading

An enzymatic mechanism by which DNA replication errors are identified and corrected.

proteome/proteomics

All the different proteins in a cell or organism/study of the same.

proximal (of a chromosomal location)

Comparatively close to the centromere.

pseudoautosomal regions (or sequences) (PAR)

Regions with identical genes at the tip of the short arms and, separately, at the tips of the long arms of the X and Y chromosomes (Figure 5.7). Because of X–Y recombination, these genes move between the X and the Y (Figure 5.8), behaving as alleles that show an apparently autosomal mode of inheritance.

pseudogene

A DNA sequence that shows a high degree of sequence homology to a non-allelic functional gene but is itself nonfunctional or does not make a protein like its closely related homolog (but it may, however, make a functional non-coding RNA) (Box 2.4).

purifying (negative) selection

A form of natural selection in which harmful mutations that wreck or disturb the function of an important DNA sequence tend to be removed from the population.

purine
A double-ringed organic nitrogenous base that is a constituent of a nucleic acid, notably adenine (A) and guanine (G)—see Figure 1.3.

pyrimidine
A single-ringed organic nitrogenous base that is a constituent of a nucleic acid, notably cytosine (C), thymine (T), and uracil (U)—see Figure 1.3.

quantitative character
A character such as height, which everybody has but to differing degrees (in contrast with a dichotomous character such as polydactyly, which some people have and others do not).

quantitative PCR (qPCR)
PCR methods that allow accurate estimation of the amount of template present (Section 3.2). See also *real-time PCR*.

quantitative trait locus (QTL)
A locus that contributes to determining the phenotype of a continuous character.

reactive oxygen species (ROS)
Chemically reactive molecules or atom containing oxygen, such as oxygen ions, oxygen radicals, and peroxides. Formed within cells as a natural by-product of normal oxygen metabolism, they have important roles in cell signaling and homeostasis but cause DNA damage (see Section 4.1).

reading frame
During translation, the way in which the continuous sequence of the mRNA is read as a series of triplet codons. There are three possible forward reading frames for any mRNA, and the correct reading frame is set by correct recognition of the AUG initiation codon (see Box 2.1).

real-time PCR
A form of quantitative PCR in which the accumulation of product is followed in real time, allowing accurate quantitation of the amount of template present (Section 3.2).

recessive
Referring to a character that is manifested in the homozygote but not in the heterozygous state.

recombinant
In linkage analysis, a gamete that contains a haplotype with a combination of alleles that is different from the combination that the parent had inherited s (Figure 8.5).

recombinant DNA
An artificially constructed hybrid DNA containing covalently linked sequences (Figure 2 in Box 3.1).

recombination (or crossover)
Exchange of DNA sequences between paired homologous chromosomes at meiosis (Figures 1.13 and 1.14).

regenerative medicine
Using stem cell cultures to provided replacement cells for cells lost through disease (or injury).

relative risk
In epidemiology, the relative risks of developing a condition in people with and without a susceptibility factor (Table 8.3).

replication fork
In DNA replication, the point along a DNA strand where the replication machinery is currently at work (Figure 1.5).

replication origin

See *origin of replication*.

replication slippage
A mistake in replication of a short tandemly repeated DNA sequence that results in newly synthesized DNA strands with more or fewer copies of the tandem repeats than in the template DNA (Figure 4.6).

reprogramming (cellular, nuclear or epigenetic)
Large-scale epigenetic changes to convert the pattern of gene expression in a cell to that typical of another cell type or cell state. Often occurs in cancers (Section 10.3) and can be artificially induced in cells (Box 9.1).

restriction endonuclease
A bacterial enzyme that cuts double-stranded DNA at a short (normally 4, 6, or 8 bp long) recognition sequence (Box 3.1).

restriction fragment length polymorphism (RFLP)
A DNA polymorphism that creates or abolishes a recognition sequence for a restriction endonuclease. When DNA is digested with the relevant enzyme, the sizes of the fragments will differ, depending on the presence or absence of the restriction site (Figure 4.4).

restriction site
A site on a DNA molecule that is cleaved by a restriction endonuclease.

retrogene
A functional gene that appears to be derived from a reverse-transcribed RNA (Box 2.4).

retroposon (or retrotransposon)
A member of a family of mobile DNA elements that transpose by making an RNA that is copied into a cDNA which integrates elsewhere in the genome (Section 2.5).

retrovirus

An RNA virus with a reverse transcriptase function, enabling the RNA genome to be copied into cDNA before integration into the chromosomes of a host cell.

reverse transcriptase
An enzyme that makes a DNA copy of an RNA template; an RNA-dependent DNA polymerase (Table 1.1).

ribozyme
A natural or synthetic catalytic RNA molecule.

risk ratio
In family studies, the relative risk of disease in a relative of an affected person compared with that of a member of the general population (Section 8.2).

RNA gene
A gene that makes a functional noncoding RNA (Figure 2.7).

RNA interference (RNAi)
A cellular defense system activated by the presence of long double-stranded RNA sequences and designed to protect against viruses and excessive transposon activity within cells (Figure 9.20). Its discovery allowed specific *gene silencing/suppression* using *siRNAs.*

RNA polymerase
An enzyme that can add ribonucleotides to the 3′ end of an RNA chain. Most RNA polymerases use a DNA template to make an RNA transcript.

RNA processing
The processes required to convert a primary transcript into a mature messenger RNA, notably capping, splicing, and polyadenylation.

RNA sequencing/RNA-Seq.

Sequencing cDNA as an indirect method of sequencing RNA. Some new technologies in principle allow direct sequencing of RNA.

RNA splicing
See *splicing*.

secondary structure
The path of the backbone of a folded polypeptide or single-stranded nucleic acid, determined by weak interactions between residues in different parts of the sequence (Box 2.2).

segmental duplication
The existence of very large, highly related DNA sequence blocks on different chromosomes, or at more than one location within a chromosome.

segregation
1. The distribution of allelic sequences between daughter cells at meiosis. Allelic sequences are said to segregate, non-allelic sequences to assort. 2. In pedigree analysis, the probability of a child's inheriting a phenotype from a parent.

selection
See *natural selection*.

selective sweep
Process whereby *positive selection* for a favorable DNA variant causes a reduction in variation in the population at the immediately neighboring nucleotide sequences (Box 4.2).

sense strand
The DNA strand of a gene that is complementary in sequence to the template (antisense) strand and identical to the transcribed RNA sequence (except that DNA contains T where RNA has U). Quoted gene

sequences always give the sense strand, in the 5′ → 3′ direction (Figure 2.1).

sensitivity (of a test)
The proportion of all true positives that the test is able to detect (Table 11.3).

sib
Brother or sister.

silencer
Combination of short DNA sequence elements that suppress the transcription of a gene.

silent mutation
Has the same meaning as *synonymous substitution*, which is the preferred term because sometimes this type of change can result in altered gene expression and disease (Figure 7.4B).

single nucleotide polymorphism/variant
See *SNP/SNV*.

siRNA (short interfering RNA)
Double-stranded RNA molecules 21–22 nucleotides long that can dramatically shut down the expression of genes through RNA interference (Figure 9.21).

sister chromatid
One of the two paired chromatids of a single chromosome that form after DNA replication and remain joined at the centromere until the anaphase stage of mitosis. Non-sister chromatids are present on different but homologous chromosomes (Figure 1.10).

SNP (single nucleotide polymorphism)

A nucleotide position in the genome where two or occasionally three alternative nucleotides are common in the population. The dbSNP database lists human SNPs but includes some rare pathogenic variants and some variants that involve two or more contiguous nucleotides.

SNV (single nucleotide variant)

A rare DNA variant (frequency less than 0.01) that can be seen to differ at a single nucleotide position from the consensus sequence in the population.

somatic cell

Any cell in the body that is not part of the germ line.

specificity (of a test for a condition)

A measure of the performance of a test that assesses the proportion of all people who do not have the condition who are correctly identified as such by the test assay. Specificity = (1 – false positive rate) (Table 11.3).

splice acceptor site

The site that defines the junction between the end of an intron in RNA and the start of the following exon. The junction sequence often conforms to the consensus sequence yyyyyyyyyyynyagR, where y is a pyrimidine, n is any nucleotide, and R is a purine that is the first nucleotide of the exon.

splice donor site

The site that defines the junction between the end of an exon in RNA and the start of the following intron. The junction sequence often conforms to the consensus sequence (C/A)AGguragu, where r is a purine and capital letters denote the end nucleotides of the exon.

splicing

The process whereby some precursor RNA transcripts are cleaved into sequences, some of which (exons) are retained and fused (spliced) to give

the mature RNA whereas others are discarded (introns).

stem cell

A cell that can act as a precursor to differentiated cells but retains the capacity for self-renewal. Can be a tissue stem cell that gives rise to a limited number of cell types (Figure 9.17 and Box and 10.2) or a *pluripotent* stem cell (Box 9.1).

stop (termination) codon

An in-frame codon that does not specify an amino acid but instead acts as a signal for the ribosome to dissociate from the mRNA and release the nascent polypeptide. See Figure 2.3 for the principle and Figure 7.2 for the different types of stop codon.

stratification

A population is stratified if it consists of several subpopulations that do not interbreed freely. Stratification is a source of error in association studies and risk estimation.

stratified medicine

A model of health care in which different medical treatments are targeted to subsets of the same disease according to which disease-associated genetic variants a person possesses.

stringency (of hybridization)

The choice of conditions that will allow either imperfectly matched sequences or only perfectly matched sequences to hybridize (Figure 3.7).

stroma

Supportive tissue of an epithelial organ, tumor and so on, consisting of connective tissues and blood vessels.

structural variation

Large-scale DNA variation that involves moving or changing the copy number of moderately long to very long DNA sequences, by one of various mechanisms: translocation, inversion, insertion, deletion, or duplication (Section 4.2).

supplementation therapy
(also called augmentation therapy) Therapy intended to supplement some deficiency, as opposed to the great majority of drug therapies that are designed to inhibit some disease process.

susceptibility factor
A variant that provides increased risk of developing a specific disease.

synonymous substitution (or silent mutation)
A nucleotide substitution that changes the sequence of a codon without any change in the amino acid that it specifies, but some may cause altered splicing and disease (Figure 7.4B).

tandem repeats
Any pattern in which a sequence of one or more nucleotides in DNA is repeated and the repetitions are directly adjacent to each other. See Figure 2.12A for an example.

targeted DNA sequencing
The process in which a defined subset of a genome (containing target sequences of interest) is captured then submitted for DNA sequencing (Box 11.2).

telomere
Specialized structure that stabilizes the ends of linear chromosomes. See Figure 1.9 for telomeric DNA structure.

termination codon
See *stop codon*.

terminal differentiation
The state of a cell that has ceased dividing and has become irreversibly committed to some specialized function.

therapeutic window
The range of plasma drug concentrations that are of therapeutic benefit without causing extra safety risks due to drug toxicity.

tissue
A set of contiguous functionally related cells.

trait
See *character*.

***trans*-acting**
The term used to describe any gene regulation in which the expression of some sequence on a DNA or RNA molecule is regulated by a different molecule or molecular assembly (in practice, a different RNA or a protein that is usually expressed from a remote gene and needs to diffuse to its site of action) (Figures 6.1, 6.2).

transcription factor
DNA-binding protein that promotes the transcription of genes. Some are ubiquitous, promoting transcription in all cells, but many are tissue-specific.

transcription unit
A segment of DNA that is used to make a primary RNA transcript (see Figure 2.1). May occasionally span multiple genes, as in the transcription of mtDNA (Figure 2.12) and in the transcription of adjacent 28S, 5.8S, and 18S rRNA genes.

transcriptome/transcriptomics
All the different RNA transcripts in a cell or tissue/the study of the same.

transdifferentiation

Epigenetic reprogramming of the nucleus of a cell, causing it to change from one cell type to another, such as from a skin cell to a neuron.

transduction

1. Relaying a signal from a cell surface receptor to a target within a cell.
2. Using recombinant viruses to introduce foreign DNA into a cell.

transfection

Direct introduction of an exogenous DNA molecule into a cell without using a vector.

transformation (of a cell)

1. Uptake by a competent microbial cell of naked high-molecular-weight DNA from the environment. 2. Alteration of the growth properties of a normal eukaryotic cell as a step toward evolving into a tumor cell.

transgene

An exogenous gene that has been transfected into cells of an animal or plant. It may be present in some tissues (as in human gene therapies) or in all tissues (as in germ-line engineering, for example in the mouse—see Box 9.2). Introduced transgenes may integrated into host cell chromosomes or replicate extrachromosomally and be transiently expressed.

transgenic animal

An animal in which artificially introduced foreign DNA (a transgene) becomes stably incorporated into the germ line (Box 9.2).

transit amplifying cells

The immediate progeny by which stem cells give rise to differentiated cells. Transit amplifying cells go through many cycles of division, but they eventually differentiate (Boxes 9.12 and 10.2).

translocation
Transfer of chromosomal regions between nonhomologous chromosomes (Figure 7.13).

transposon/transposon repeat
A mobile genetic element/a member of a repetitive DNA family containing some members that are able to transpose but also many inactivated copies of transposons (Table 2.6, Figure 2.15).

trophoblast
Outer layer of polarized cells in the blastocyst that will go on to form the chorion, the embryonic component of the placenta (Figure 2 of Box 9.1, Figure 2).

tropism
The specificity of a virus for a particular cell type, determined in part by the interaction of viral surface structures with receptors present on the surface of the cell.

tumor suppressor gene
A gene that is commonly inactivated in tumors (by an inactivating mutation, by deletion as a result of abnormal chromosome segregation/recombination, or by epigenetic silencing). Classical tumor suppressor genes normally work to inhibit or control cell division.

unequal crossover
Recombination between chromatids that have paired up slightly out of alignment. See Figure 7.8.

unequal sister chromatid exchange
The same process as *unequal crossover* but involves sister chromatids. See Figure 7.8.

uniparental diploidy

A 46,XX diploid conceptus in which both genomes derive from the same parent. Such conceptuses never develop normally (Figure 6.19).

uniparental disomy
A cell or organism in which both copies of one particular chromosome pair are derived from one parent. Depending on the chromosome involved, this may or may not cause disease (Figure 6.24).

unrelated
Ultimately everybody is related; the word is used in this book to mean people who do not have an identified common ancestor in the last four or so generations.

untranslated region (5′ UTR, 3′ UTR)
Regions at the 5′ end of mRNA before the AUG translation start codon, or at the 3′ end after the stop codon (Figures 2.1, and 2.3).

variant (in relation to DNA)
A sequence that is different from the majority sequence but exists at a low frequency (<0.01; that is, less than 1%) in the population.

vector
A nucleic acid that is able to replicate and maintain itself within a host cell and that can be used to confer similar properties on any sequence covalently linked to it.

X-chromosome inactivation (or X-inactivation)
The *epigenetic* inactivation of all except one of the X chromosomes in the cells of humans and other mammals that have more than one X (Figure 6.20).

zygote
The fertilized egg cell.

Index

Note: The index covers the main text but not the summaries, questions or glossary.

Prefixes have been ignored for filing unless integral to a topic (so ‘β-sheets’ at ‘beta’ but ‘β-thalassemia’ at ‘thalassemia’). The same apples to numeric prefixes but unavoidable numerics have been sorted as though spelled out (so ‘5-methlycytosine’ will be at ‘methyl’ but ‘7SL RNA’ at ‘seven’).

Page numbers ending with ‘B’, ‘F’ or ‘T’ indicate that the listed topic is dealt with on that page *only* in a box, figure or table. ff. following a page number means that the item appears not only in the given page number but also in multiple consecutive pages.

A

B

C

D

G

H

I

J

K

L

O

P

Q

R

S

T

U

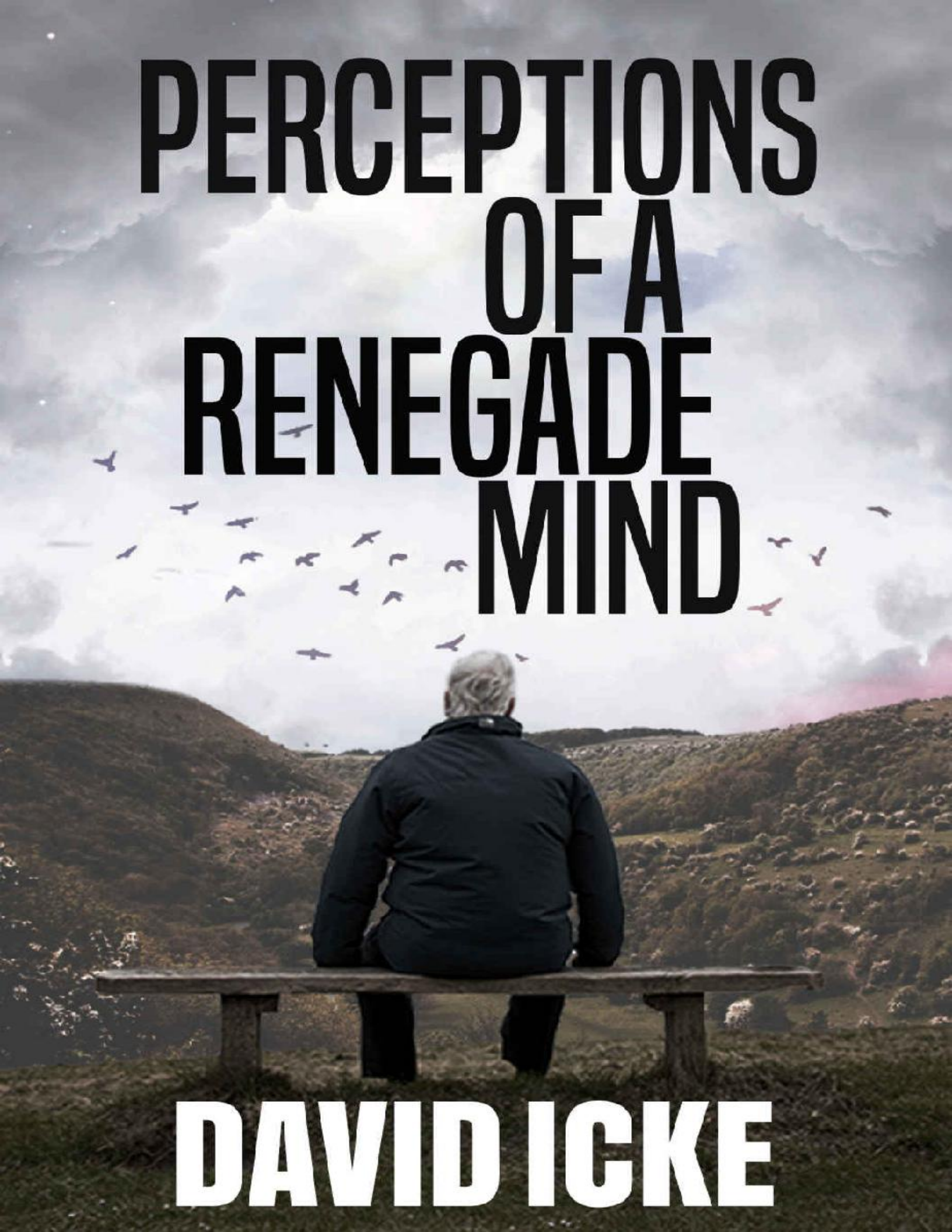
W

X

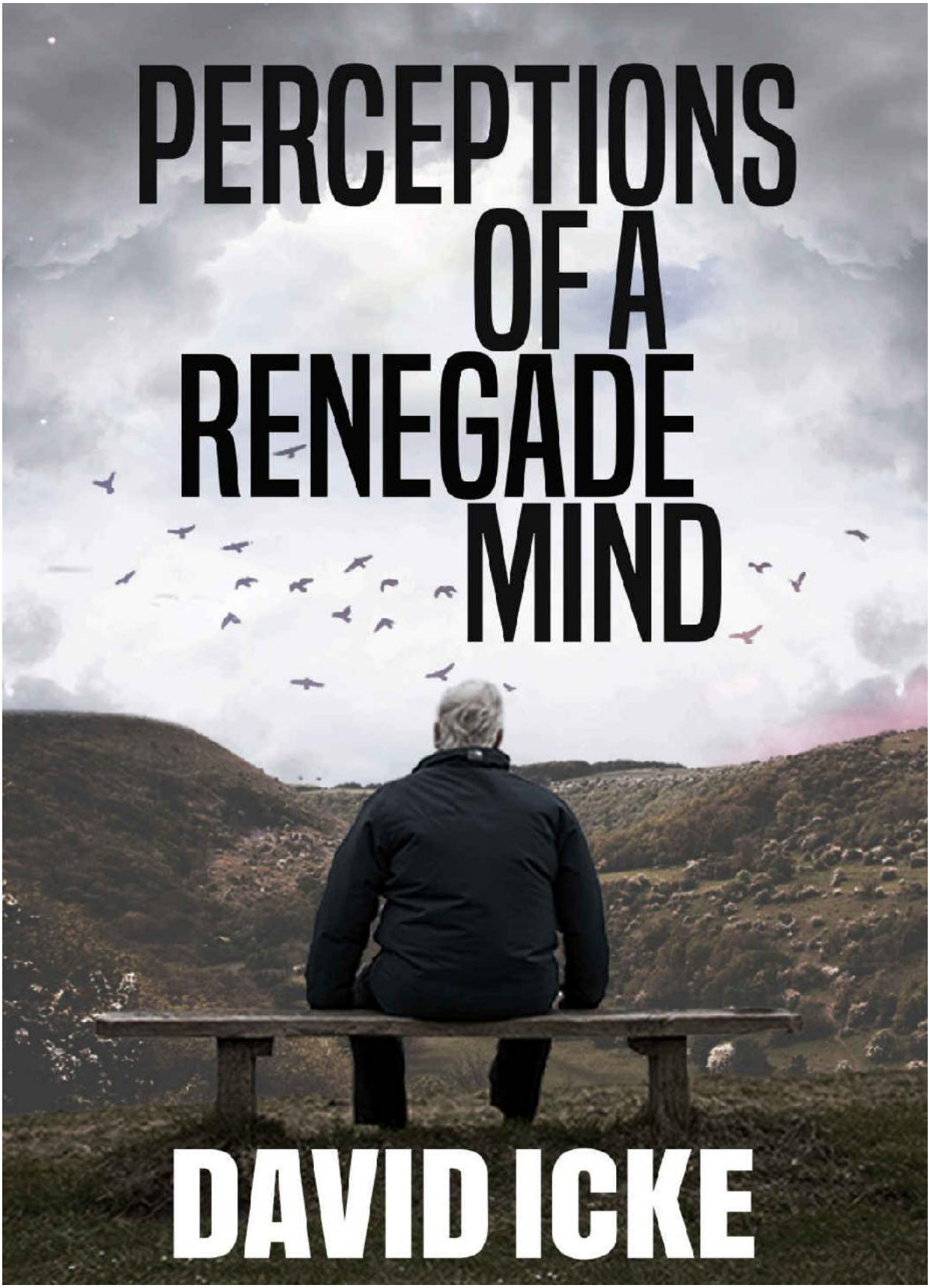
Y

Z

PERCEPTIONS OF A RENEGADE MIND



DAVID ICKE


PERCEPTIONS
OF A
RENEGADE
MIND
DAVID ICKE

PERCEPTIONS OF A RENEGADE MIND

ickonic
publishing

First published in July 2021.

ickonic
publishing

New Enterprise House
St Helens Street
Derby
DE1 3GY
UK

email: gareth.icke@davidicke.com

Cover Design: Gareth Icke
Book Design: Neil Hague

British Library Cataloguing-in Publication Data
A catalogue record for this book is available from the British Library

eISBN 978-18384153-1-0

PERCEPTIONS OF A RENEGADE MIND

DAVID ICKE

Dedication:

To *Freeeeeedom!*

Renegade:
Adjective
‘Having rejected tradition: Unconventional.’
Merriam-Webster Dictionary

Acquiescence to tyranny is the death of the spirit

You may be 38 years old, as I happen to be. And one day, some great opportunity stands before you and calls you to stand up for some great principle, some great issue, some great cause. And you refuse to do it because you are afraid … You refuse to do it because you want to live longer … You're afraid that you will lose your job, or you are afraid that you will be criticised or that you will lose your popularity, or you're afraid that somebody will stab you, or shoot at you or bomb your house; so you refuse to take the stand.

Well, you may go on and live until you are 90, but you're just as dead at 38 as you would be at 90. And the cessation of breathing in your life is but the belated announcement of an earlier death of the spirit.

Martin Luther King

How the few control the many and always have – the many do whatever they're told

'Forward, the Light Brigade!'
Was there a man dismayed?
Not though the soldier knew
Someone had blundered.
Theirs not to make reply,
Theirs not to reason why,
Theirs but to do and die.
Into the valley of Death
Rode the six hundred.

Cannon to right of them,
Cannon to left of them,
Cannon in front of them
Volleyed and thundered;
Stormed at with shot and shell,
Boldly they rode and well,
Into the jaws of Death,
Into the mouth of hell
Rode the six hundred

Alfred Lord Tennyson (1809-1892)

The mist is lifting slowly
I can see the way ahead
And I've left behind the empty streets
That once inspired my life
And the strength of the emotion
Is like thunder in the air
'Cos the promise that we made each other
Haunts me to the end

The secret of your beauty
And the mystery of your soul
I've been searching for in everyone I meet
And the times I've been mistaken
It's impossible to say
And the grass is growing
Underneath our feet

The words that I remember
From my childhood still are true
That there's none so blind
As those who will not see
And to those who lack the courage
And say it's dangerous to try
Well they just don't know
That love eternal will not be denied

I know you're out there somewhere
Somewhere, somewhere
I know you're out there somewhere

Somewhere you can hear my voice
I know I'll find you somehow
Somehow, somehow
I know I'll find you somehow
And somehow I'll return again to you

The Moody Blues

Are you a gutless wonder - or a Renegade Mind?

Monuments put from pen to paper,
Turns me into a gutless wonder,
And if you tolerate this,
Then your children will be next.
Gravity keeps my head down,
Or is it maybe shame ...

Manic Street Preachers

Rise like lions after slumber
In unvanquishable number.
Shake your chains to earth like dew
Which in sleep have fallen on you.
Ye are many – they are few.

Percy Shelley

Contents

CHAPTER ONE

I'm thinking' – Oh, but *are* you?

Think for yourself and let others enjoy the privilege of doing so too
Voltaire

French-born philosopher, mathematician and scientist René Descartes became famous for his statement in Latin in the 17th century which translates into English as: 'I think, therefore I am.'

On the face of it that is true. Thought reflects perception and perception leads to both behaviour and self-identity. In that sense 'we' are what we think. But who or what is doing the thinking and is thinking the only route to perception? Clearly, as we shall see, 'we' are not always the source of 'our' perception, indeed with regard to humanity as a whole this is rarely the case; and thinking is far from the only means of perception. Thought is the village idiot compared with other expressions of consciousness that we all have the potential to access and tap into. This has to be true when we *are* those other expressions of consciousness which are infinite in nature. We have forgotten this, or, more to the point, been manipulated to forget.

These are not just the esoteric musings of the navel. The whole foundation of human control and oppression is control of perception. Once perception is hijacked then so is behaviour which is dictated by perception. Collective perception becomes collective behaviour and collective behaviour is what we call human society. Perception is all and those behind human control know that which is

why perception is the target 24/7 of the psychopathic manipulators that I call the Global Cult. They know that if they dictate perception they will dictate behaviour and collectively dictate the nature of human society. They are further aware that perception is formed from information received and if they control the circulation of information they will to a vast extent direct human behaviour. Censorship of information and opinion has become globally Nazi-like in recent years and never more blatantly than since the illusory 'virus pandemic' was triggered out of China in 2019 and across the world in 2020. Why have billions submitted to house arrest and accepted fascistic societies in a way they would have never believed possible? Those controlling the information spewing from government, mainstream media and Silicon Valley (all controlled by the same Global Cult networks) told them they were in danger from a 'deadly virus' and only by submitting to house arrest and conceding their most basic of freedoms could they and their families be protected. This monumental and provable lie became the *perception* of the billions and therefore the *behaviour* of the billions. In those few words you have the whole structure and modus operandi of human control. Fear is a perception – **F**alse **E**motion **A**ppearing **R**eal – and fear is the currency of control. In short … get them by the balls (or give them the impression that you have) and their hearts and minds will follow. Nothing grips the dangly bits and freezes the rear-end more comprehensively than fear.

World number 1

There are two 'worlds' in what appears to be one 'world' and the prime difference between them is knowledge. First we have the mass of human society in which the population is maintained in coldly-calculated ignorance through control of information and the 'education' (indoctrination) system. That's all you really need to control to enslave billions in a perceptual delusion in which what are perceived to be *their* thoughts and opinions are ever-repeated mantras that the system has been downloading all their lives through 'education', media, science, medicine, politics and academia

in which the personnel and advocates are themselves overwhelmingly the perceptual products of the same repetition. Teachers and academics in general are processed by the same programming machine as everyone else, but unlike the great majority they never leave the 'education' program. It gripped them as students and continues to grip them as programmers of subsequent generations of students. The programmed become the programmers – the programmed programmers. The same can largely be said for scientists, doctors and politicians and not least because as the American writer Upton Sinclair said: 'It is difficult to get a man to understand something when his salary depends upon his not understanding it.' If your career and income depend on thinking the way the system demands then you will – bar a few free-minded exceptions – concede your mind to the Perceptual Mainframe that I call the Postage Stamp Consensus. This is a tiny band of perceived knowledge and possibility 'taught' (downloaded) in the schools and universities, pounded out by the mainstream media and on which all government policy is founded. Try thinking, and especially speaking and acting, outside of the 'box' of consensus and see what that does for your career in the Mainstream Everything which bullies, harasses, intimidates and ridicules the population into compliance. Here we have the simple structure which enslaves most of humanity in a perceptual prison cell for an entire lifetime and I'll go deeper into this process shortly. Most of what humanity is taught as fact is nothing more than programmed belief. American science fiction author Frank Herbert was right when he said: 'Belief can be manipulated. Only knowledge is dangerous.' In the 'Covid' age belief is promoted and knowledge is censored. It was always so, but never to the extreme of today.

World number 2

A 'number 2' is slang for 'doing a poo' and how appropriate that is when this other 'world' is doing just that on humanity every minute of every day. World number 2 is a global network of secret societies and semi-secret groups dictating the direction of society via

governments, corporations and authorities of every kind. I have spent more than 30 years uncovering and exposing this network that I call the Global Cult and knowing its agenda is what has made my books so accurate in predicting current and past events. Secret societies are secret for a reason. They want to keep their hoarded knowledge to themselves and their chosen initiates and to hide it from the population which they seek through ignorance to control and subdue. The whole foundation of the division between World 1 and World 2 is *knowledge*. What number 1 knows number 2 must not. Knowledge they have worked so hard to keep secret includes (a) the agenda to enslave humanity in a centrally-controlled global dictatorship, and (b) the nature of reality and life itself. The latter (b) must be suppressed to allow the former (a) to prevail as I shall be explaining. The way the Cult manipulates and interacts with the population can be likened to a spider's web. The 'spider' sits at the centre in the shadows and imposes its will through the web with each strand represented in World number 2 by a secret society, satanic or semi-secret group, and in World number 1 – the world of the seen – by governments, agencies of government, law enforcement, corporations, the banking system, media conglomerates and Silicon Valley (Fig 1 overleaf). The spider and the web connect and coordinate all these organisations to pursue the same global outcome while the population sees them as individual entities working randomly and independently. At the level of the web governments *are* the banking system *are* the corporations *are* the media *are* Silicon Valley *are* the World Health Organization working from their inner cores as one unit. Apparently unconnected countries, corporations, institutions, organisations and people are on the *same team* pursuing the same global outcome. Strands in the web immediately around the spider are the most secretive and exclusive secret societies and their membership is emphatically restricted to the Cult inner-circle emerging through the generations from particular bloodlines for reasons I will come to. At the core of the core you would get them in a single room. That's how many people are dictating the direction of human society and its transformation

through the ‘Covid’ hoax and other means. As the web expands out from the spider we meet the secret societies that many people will be aware of – the Freemasons, Knights Templar, Knights of Malta, Opus Dei, the inner sanctum of the Jesuit Order, and such like. Note how many are connected to the Church of Rome and there is a reason for that. The Roman Church was established as a revamp, a rebranding, of the relocated ‘Church’ of Babylon and the Cult imposing global tyranny today can be tracked back to Babylon and Sumer in what is now Iraq.

Figure 1: The global web through which the few control the many. (Image Neil Hague.)

Inner levels of the web operate in the unseen away from the public eye and then we have what I call the cusp organisations located at the point where the hidden meets the seen. They include a series of satellite organisations answering to a secret society founded in London in the late 19th century called the Round Table and among them are the Royal Institute of International Affairs (UK, founded in 1920); Council on Foreign Relations (US, 1921); Bilderberg Group (worldwide, 1954); Trilateral Commission (US/worldwide, 1972); and the Club of Rome (worldwide, 1968) which was created to exploit environmental concerns to justify the centralisation of global power to ‘save the planet’. The Club of Rome instigated with others the human-caused climate change hoax which has led to all the ‘green

new deals' demanding that very centralisation of control. Cusp organisations, which include endless 'think tanks' all over the world, are designed to coordinate a single global policy between political and business leaders, intelligence personnel, media organisations and anyone who can influence the direction of policy in their own sphere of operation. Major players and regular attenders will know what is happening – or some of it – while others come and go and are kept overwhelmingly in the dark about the big picture. I refer to these cusp groupings as semi-secret in that they can be publicly identified, but what goes on at the inner-core is kept very much 'in house' even from most of their members and participants through a fiercely-imposed system of compartmentalisation. Only let them know what they need to know to serve your interests and no more. The structure of secret societies serves as a perfect example of this principle. Most Freemasons never get higher than the bottom three levels of 'degree' (degree of knowledge) when there are 33 official degrees of the Scottish Rite. Initiates only qualify for the next higher 'compartment' or degree if those at that level choose to allow them. Knowledge can be carefully assigned only to those considered 'safe'. I went to my local Freemason's lodge a few years ago when they were having an 'open day' to show how cuddly they were and when I chatted to some of them I was astonished at how little the rank and file knew even about the most ubiquitous symbols they use. The mushroom technique – keep them in the dark and feed them bullshit – applies to most people in the web as well as the population as a whole. Sub-divisions of the web mirror in theme and structure transnational corporations which have a headquarters somewhere in the world dictating to all their subsidiaries in different countries. Subsidiaries operate in their methodology and branding to the same centrally-dictated plan and policy in pursuit of particular ends. The Cult web functions in the same way. Each country has its own web as a subsidiary of the global one. They consist of networks of secret societies, semi-secret groups and bloodline families and their job is to impose the will of the spider and the global web in their particular country. Subsidiary networks control and manipulate the national political system, finance, corporations, media, medicine, etc. to

ensure that they follow the globally-dictated Cult agenda. These networks were the means through which the 'Covid' hoax could be played out with almost every country responding in the same way.

The 'Yessir' pyramid

Compartmentalisation is the key to understanding how a tiny few can dictate the lives of billions when combined with a top-down sequence of imposition and acquiescence. The inner core of the Cult sits at the peak of the pyramidal hierarchy of human society (Fig 2 overleaf). It imposes its will – its agenda for the world – on the level immediately below which acquiesces to that imposition. This level then imposes the Cult will on the level below them which acquiesces and imposes on the next level. Very quickly we meet levels in the hierarchy that have no idea there even is a Cult, but the sequence of imposition and acquiescence continues down the pyramid in just the same way. 'I don't know why we are doing this but the order came from "on-high" and so we better just do it.' Alfred Lord Tennyson said of the cannon fodder levels in his poem *The Charge of the Light Brigade*: 'Theirs not to reason why; theirs but to do and die.' The next line says that 'into the valley of death rode the six hundred' and they died because they obeyed without question what their perceived 'superiors' told them to do. In the same way the population capitulated to 'Covid'. The whole hierarchical pyramid functions like this to allow the very few to direct the enormous many. Eventually imposition-acquiescence-imposition-acquiescence comes down to the mass of the population at the foot of the pyramid. If they acquiesce to those levels of the hierarchy imposing on them (governments/law enforcement/doctors/media) a circuit is completed between the population and the handful of super-psychopaths in the Cult inner core at the top of the pyramid. Without a circuit-breaking refusal to obey, the sequence of imposition and acquiescence allows a staggeringly few people to impose their will upon the entirety of humankind. We are looking at the very sequence that has subjugated billions since the start of 2020. Our freedom has not been taken from us. Humanity has given it

away. Fascists do not impose fascism because there are not enough of them. Fascism is imposed by the population acquiescing to fascism. Put another way allowing their perceptions to be programmed to the extent that leads to the population giving their freedom away by giving their perceptions – their mind – away. If this circuit is not broken by humanity ceasing to cooperate with their own enslavement then nothing can change. For that to happen people have to critically think and see through the lies and window dressing and then summon the backbone to act upon what they see. The Cult spends its days working to stop either happening and its methodology is systematic and highly detailed, but it can be overcome and that is what this book is all about.

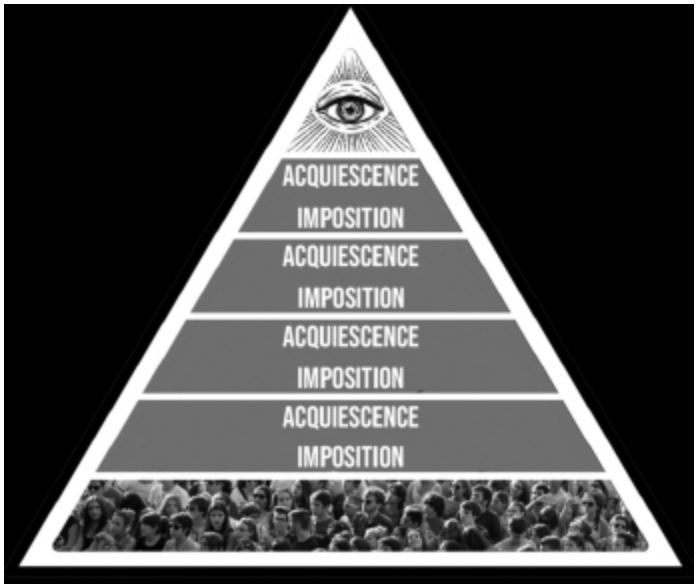


Figure 2: The simple sequence of imposition and compliance that allows a handful of people at the peak of the pyramid to dictate the lives of billions.

The Life Program

Okay, back to world number 1 or the world of the 'masses'. Observe the process of what we call 'life' and it is a perceptual download from cradle to grave. The Cult has created a global structure in which perception can be programmed and the program continually topped-up with what appears to be constant confirmation that the program is indeed true reality. The important word here is 'appears'.

This is the structure, the fly-trap, the Postage Stamp Consensus or Perceptual Mainframe, which represents that incredibly narrow band of perceived possibility delivered by the 'education' system, mainstream media, science and medicine. From the earliest age the download begins with parents who have themselves succumbed to the very programming their children are about to go through. Most parents don't do this out of malevolence and mostly it is quite the opposite. They do what they believe is best for their children and that is what the program has told them is best. Within three or four years comes the major transition from parental programming to full-blown state (Cult) programming in school, college and university where perceptually-programmed teachers and academics pass on their programming to the next generations. Teachers who resist are soon marginalised and their careers ended while children who resist are called a problem child for whom Ritalin may need to be prescribed. A few years after entering the 'world' children are under the control of authority figures representing the state telling them when they have to be there, when they can leave and when they can speak, eat, even go to the toilet. This is calculated preparation for a lifetime of obeying authority in all its forms. Reflex-action fear of authority is instilled by authority from the start. Children soon learn the carrot and stick consequences of obeying or defying authority which is underpinned daily for the rest of their life. Fortunately I daydreamed through this crap and never obeyed authority simply because it told me to. This approach to my alleged 'betters' continues to this day. There can be consequences of pursuing open-minded freedom in a world of closed-minded conformity. I spent a lot of time in school corridors after being ejected from the classroom for not taking some of it seriously and now I spend a lot of time being ejected from Facebook, YouTube and Twitter. But I can tell you that being true to yourself and not compromising your self-respect is far more exhilarating than bowing to authority for authority's sake. You don't have to be a sheep to the shepherd (authority) and the sheep dog (fear of not obeying authority).

The perceptual download continues throughout the formative years in school, college and university while script-reading 'teachers', 'academics' 'scientists', 'doctors' and 'journalists' insist that ongoing generations must be as programmed as they are. Accept the program or you will not pass your 'exams' which confirm your 'degree' of programming. It is tragic to think that many parents pressure their offspring to work hard at school to download the program and qualify for the next stage at college and university. The late, great, American comedian George Carlin said: 'Here's a bumper sticker I'd like to see: We are proud parents of a child who has resisted his teachers' attempts to break his spirit and bend him to the will of his corporate masters.' Well, the best of luck finding many of those, George. Then comes the moment to leave the formal programming years in academia and enter the 'adult' world of work. There you meet others in your chosen or prescribed arena who went through the same Postage Stamp Consensus program before you did. There is therefore overwhelming agreement between almost everyone on the basic foundations of Postage Stamp reality and the rejection, even contempt, of the few who have a mind of their own and are prepared to use it. This has two major effects. Firstly, the consensus confirms to the programmed that their download is really how things are. I mean, everyone knows that, right? Secondly, the arrogance and ignorance of Postage Stamp adherents ensure that anyone questioning the program will have unpleasant consequences for seeking their own truth and not picking their perceptions from the shelf marked: 'Things you must believe without question and if you don't you're a dangerous lunatic conspiracy theorist and a harebrained nutter'.

Every government, agency and corporation is founded on the same Postage Stamp prison cell and you can see why so many people believe the same thing while calling it their own 'opinion'. Fusion of governments and corporations in pursuit of the same agenda was the definition of fascism described by Italian dictator Benito Mussolini. The pressure to conform to perceptual norms downloaded for a lifetime is incessant and infiltrates society right

down to family groups that become censors and condemners of their own 'black sheep' for not, ironically, being sheep. We have seen an explosion of that in the 'Covid' era. Cult-owned global media unleashes its propaganda all day every day in support of the Postage Stamp and targets with abuse and ridicule anyone in the public eye who won't bend their mind to the will of the tyranny. Any response to this is denied (certainly in my case). They don't want to give a platform to expose official lies. Cult-owned-and-created Internet giants like Facebook, Google, YouTube and Twitter delete you for having an unapproved opinion. Facebook boasts that its AI censors delete 97-percent of 'hate speech' before anyone even reports it. Much of that 'hate speech' will simply be an opinion that Facebook and its masters don't want people to see. Such perceptual oppression is widely known as fascism. Even Facebook executive Benny Thomas, a 'CEO Global Planning Lead', said in comments secretly recorded by investigative journalism operation Project Veritas that Facebook is 'too powerful' and should be broken up:

> I mean, no king in history has been the ruler of two billion people, but Mark Zuckerberg is … And he's 36. That's too much for a 36-year-old ... You should not have power over two billion people. I just think that's wrong.

Thomas said Facebook-owned platforms like Instagram, Oculus, and WhatsApp needed to be separate companies. 'It's too much power when they're all one together'. That's the way the Cult likes it, however. We have an executive of a Cult organisation in Benny Thomas that doesn't know there is a Cult such is the compartmentalisation. Thomas said that Facebook and Google 'are no longer companies, they're countries'. Actually they are more powerful than countries on the basis that if you control information you control perception and control human society.

I love my oppressor

Another expression of this psychological trickery is for those who realise they are being pressured into compliance to eventually

convince themselves to believe the official narratives to protect their self-respect from accepting the truth that they have succumbed to meek and subservient compliance. Such people become some of the most vehement defenders of the system. You can see them everywhere screaming abuse at those who prefer to think for themselves and by doing so reminding the compliers of their own capitulation to conformity. 'You are talking dangerous nonsense you Covidiot!!' Are you trying to convince me or yourself? It is a potent form of Stockholm syndrome which is defined as: 'A psychological condition that occurs when a victim of abuse identifies and attaches, or bonds, positively with their abuser.' An example is hostages bonding and even 'falling in love' with their kidnappers. The syndrome has been observed in domestic violence, abused children, concentration camp inmates, prisoners of war and many and various Satanic cults. These are some traits of Stockholm syndrome listed at goodtherapy.org:

- Positive regard towards perpetrators of abuse or captor [see 'Covid'].
- Failure to cooperate with police and other government authorities when it comes to holding perpetrators of abuse or kidnapping accountable [or in the case of 'Covid' cooperating with the police to enforce and defend their captors' demands].
- Little or no effort to escape [see 'Covid'].
- Belief in the goodness of the perpetrators or kidnappers [see 'Covid'].
- Appeasement of captors. This is a manipulative strategy for maintaining one's safety. As victims get rewarded – perhaps with less abuse or even with life itself – their appeasing behaviours are reinforced [see 'Covid'].
- Learned helplessness. This can be akin to 'if you can't beat 'em, join 'em'. As the victims fail to escape the abuse or captivity, they may start giving up and soon realize it's just easier for everyone if they acquiesce all their power to their captors [see 'Covid'].

- Feelings of pity toward the abusers, believing they are actually victims themselves. Because of this, victims may go on a crusade or mission to 'save' [protect] their abuser [see the venom unleashed on those challenging the official 'Covid' narrative].
- Unwillingness to learn to detach from their perpetrators and heal. In essence, victims may tend to be less loyal to themselves than to their abuser [*definitely* see 'Covid'].

Ponder on those traits and compare them with the behaviour of great swathes of the global population who have defended governments and authorities which have spent every minute destroying their lives and livelihoods and those of their children and grandchildren since early 2020 with fascistic lockdowns, house arrest and employment deletion to 'protect' them from a 'deadly virus' that their abusers' perceptually created to bring about this very outcome. We are looking at mass Stockholm syndrome. All those that agree to concede their freedom will believe those perceptions are originating in their own independent 'mind' when in fact by conceding their reality to Stockholm syndrome they have by definition conceded any independence of mind. Listen to the 'opinions' of the acquiescing masses in this 'Covid' era and what gushes forth is the repetition of the official version of everything delivered unprocessed, unfiltered and unquestioned. The whole programming dynamic works this way. I must be free because I'm told that I am and so I think that I am.

You can see what I mean with the chapter theme of 'I'm thinking – Oh, but *are* you?' The great majority are not thinking, let alone for themselves. They are repeating what authority has told them to believe which allows them to be controlled. Weaving through this mentality is the fear that the 'conspiracy theorists' are right and this again explains the often hysterical abuse that ensues when you dare to contest the official narrative of anything. Denial is the mechanism of hiding from yourself what you don't want to be true. Telling people what they want to hear is easy, but it's an infinitely greater challenge to tell them what they would rather not be happening.

One is akin to pushing against an open door while the other is met with vehement resistance no matter what the scale of evidence. I don't want it to be true so I'll convince myself that it's not. Examples are everywhere from the denial that a partner is cheating despite all the signs to the reflex-action rejection of any idea that world events in which country after country act in exactly the same way are centrally coordinated. To accept the latter is to accept that a force of unspeakable evil is working to destroy your life and the lives of your children with nothing too horrific to achieve that end. Who the heck wants that to be true? But if we don't face reality the end is duly achieved and the consequences are far worse and ongoing than breaking through the walls of denial today with the courage to make a stand against tyranny.

Connect the dots – but how?

A crucial aspect of perceptual programming is to portray a world in which everything is random and almost nothing is connected to anything else. Randomness cannot be coordinated by its very nature and once you perceive events as random the idea they could be connected is waved away as the rantings of the tinfoil-hat brigade. You can't plan and coordinate random you idiot! No, you can't, but you can hide the coldly-calculated and long-planned behind the *illusion* of randomness. A foundation manifestation of the Renegade Mind is to scan reality for patterns that connect the apparently random and turn pixels and dots into pictures. This is the way I work and have done so for more than 30 years. You look for similarities in people, modus operandi and desired outcomes and slowly, then ever quicker, the picture forms. For instance: There would seem to be no connection between the 'Covid pandemic' hoax and the human-caused global-warming hoax and yet they are masks (appropriately) on the same face seeking the same outcome. Those pushing the global warming myth through the Club of Rome and other Cult agencies are driving the lies about 'Covid' – Bill Gates is an obvious one, but they are endless. Why would the same people be involved in both when they are clearly not connected? Oh, but they

are. Common themes with personnel are matched by common goals. The 'solutions' to both 'problems' are centralisation of global power to impose the will of the few on the many to 'save' humanity from 'Covid' and save the planet from an 'existential threat' (we need 'zero Covid' and 'zero carbon emissions'). These, in turn, connect with the 'dot' of globalisation which was coined to describe the centralisation of global power in every area of life through incessant political and corporate expansion, trading blocks and superstates like the European Union. If you are the few and you want to control the many you have to centralise power and decision-making. The more you centralise power the more power the few at the centre will have over the many; and the more that power is centralised the more power those at the centre have to centralise even quicker. The momentum of centralisation gets faster and faster which is exactly the process we have witnessed. In this way the hoaxed 'pandemic' and the fakery of human-caused global warming serve the interests of globalisation and the seizure of global power in the hands of the Cult inner-circle which is behind 'Covid', 'climate change' *and* globalisation. At this point random 'dots' become a clear and obvious picture or pattern.

Klaus Schwab, the classic Bond villain who founded the Cult's Gates-funded World Economic Forum, published a book in 2020, *The Great Reset*, in which he used the 'problem' of 'Covid' to justify a total transformation of human society to 'save' humanity from 'climate change'. Schwab said: 'The pandemic represents a rare but narrow window of opportunity to reflect, reimagine, and reset our world.' What he didn't mention is that the Cult he serves is behind both hoaxes as I show in my book *The Answer*. He and the Cult don't have to reimagine the world. They know precisely what they want and that's why they destroyed human society with 'Covid' to 'build back better' in their grand design. Their job is not to imagine, but to get humanity to imagine and agree with their plans while believing it's all random. It must be pure coincidence that 'The Great Reset' has long been the Cult's code name for the global imposition of fascism and replaced previous code-names of the 'New World

Order' used by Cult frontmen like Father George Bush and the 'New Order of the Ages' which emerged from Freemasonry and much older secret societies. New Order of the Ages appears on the reverse of the Great Seal of the United States as 'Novus ordo seclorum' underneath the Cult symbol used since way back of the pyramid and all seeing-eye (Fig 3). The pyramid is the hierarchy of human control headed by the illuminated eye that symbolises the force behind the Cult which I will expose in later chapters. The term 'Annuit Coeptis' translates as 'He favours our undertaking'. We are told the 'He' is the Christian god, but 'He' is not as I will be explaining.

Figure 3: The all-seeing eye of the Cult 'god' on the Freemason-designed Great Seal of the United States and also on the dollar bill.

Having you on

Two major Cult techniques of perceptual manipulation that relate to all this are what I have called since the 1990s Problem-Reaction-Solution (PRS) and the Totalitarian Tiptoe (TT). They can be uncovered by the inquiring mind with a simple question: Who benefits? The answer usually identifies the perpetrators of a given action or happening through the concept of 'he who most benefits from a crime is the one most likely to have committed it'. The Latin 'Cue bono?' – Who benefits? – is widely attributed to the Roman orator and statesman Marcus Tullius Cicero. No wonder it goes back so far when the concept has been relevant to human behaviour since

history was recorded. Problem-Reaction-Solution is the technique used to manipulate us every day by covertly creating a problem (or the illusion of one) and offering the solution to the problem (or the illusion of one). In the first phase you create the problem and blame someone or something else for why it has happened. This may relate to a financial collapse, terrorist attack, war, global warming or pandemic, anything in fact that will allow you to impose the 'solution' to change society in the way you desire at that time. The 'problem' doesn't have to be real. PRS is manipulation of perception and all you need is the population to believe the problem is real. Human-caused global warming and the 'Covid pandemic' only have to be *perceived* to be real for the population to accept the 'solutions' of authority. I refer to this technique as NO-Problem-Reaction-Solution. Billions did not meekly accept house arrest from early 2020 because there was a real deadly 'Covid pandemic' but because they perceived – believed – that to be the case. The antidote to Problem-Reaction-Solution is to ask who benefits from the proposed solution. Invariably it will be anyone who wants to justify more control through deletion of freedom and centralisation of power and decision-making.

The two world wars were Problem-Reaction-Solutions that transformed and realigned global society. Both were manipulated into being by the Cult as I have detailed in books since the mid-1990s. They dramatically centralised global power, especially World War Two, which led to the United Nations and other global bodies thanks to the overt and covert manipulations of the Rockefeller family and other Cult bloodlines like the Rothschilds. The UN is a stalking horse for full-blown world government that I will come to shortly. The land on which the UN building stands in New York was donated by the Rockefellers and the same Cult family was behind Big Pharma scalpel and drug 'medicine' and the creation of the World Health Organization as part of the UN. They have been stalwarts of the eugenics movement and funded Hitler's race-purity expert' Ernst Rudin. The human-caused global warming hoax has been orchestrated by the Club of Rome through the UN which is

manufacturing both the 'problem' through its Intergovernmental Panel on Climate Change and imposing the 'solution' through its Agenda 21 and Agenda 2030 which demand the total centralisation of global power to 'save the world' from a climate hoax the United Nations is itself perpetrating. What a small world the Cult can be seen to be particularly among the inner circles. The bedfellow of Problem-Reaction-Solution is the Totalitarian Tiptoe which became the Totalitarian Sprint in 2020. The technique is fashioned to hide the carefully-coordinated behind the cover of apparently random events. You start the sequence at 'A' and you know you are heading for 'Z'. You don't want people to know that and each step on the journey is presented as a random happening while all the steps strung together lead in the same direction. The speed may have quickened dramatically in recent times, but you can still see the incremental approach of the Tiptoe in the case of 'Covid' as each new imposition takes us deeper into fascism. Tell people they have to do this or that to get back to 'normal', then this and this and this. With each new demand adding to the ones that went before the population's freedom is deleted until it disappears. The spider wraps its web around the flies more comprehensively with each new diktat. I'll highlight this in more detail when I get to the 'Covid' hoax and how it has been pulled off. Another prime example of the Totalitarian Tiptoe is how the Cult-created European Union went from a 'free-trade zone' to a centralised bureaucratic dictatorship through the Tiptoe of incremental centralisation of power until nations became mere administrative units for Cult-owned dark suits in Brussels.

The antidote to ignorance is knowledge which the Cult seeks vehemently to deny us, but despite the systematic censorship to that end the Renegade Mind can overcome this by vociferously seeking out the facts no matter the impediments put in the way. There is also a method of thinking and perceiving – *knowing* – that doesn't even need names, dates, place-type facts to identify the patterns that reveal the story. I'll get to that in the final chapter. All you need to know about the manipulation of human society and to what end is still out there – *at the time of writing* – in the form of books, videos

and websites for those that really want to breach the walls of programmed perception. To access this knowledge requires the abandonment of the mainstream media as a source of information in the awareness that this is owned and controlled by the Cult and therefore promotes mass perceptions that suit the Cult. Mainstream media lies all day, every day. That is its function and very reason for being. Where it does tell the truth, here and there, is only because the truth and the Cult agenda very occasionally coincide. If you look for fact and insight to the BBC, CNN and virtually all the rest of them you are asking to be conned and perceptually programmed.

Know the outcome and you'll see the journey

Events seem random when you have no idea where the world is being taken. Once you do the random becomes the carefully planned. Know the outcome and you'll see the journey is a phrase I have been using for a long time to give context to daily happenings that appear unconnected. Does a problem, or illusion of a problem, trigger a proposed 'solution' that further drives society in the direction of the outcome? Invariably the answer will be yes and the random – *abracadabra* – becomes the clearly coordinated. So what is this outcome that unlocks the door to a massively expanded understanding of daily events? I will summarise its major aspects – the fine detail is in my other books – and those new to this information will see that the world they thought they were living in is a very different place. The foundation of the Cult agenda is the incessant centralisation of power and all such centralisation is ultimately in pursuit of Cult control on a global level. I have described for a long time the planned world structure of top-down dictatorship as the Hunger Games Society. The term obviously comes from the movie series which portrayed a world in which a few living in military-protected hi-tech luxury were the overlords of a population condemned to abject poverty in isolated 'sectors' that were not allowed to interact. 'Covid' lockdowns and travel bans anyone? The 'Hunger Games' pyramid of structural control has the inner circle of the Cult at the top with pretty much the entire

population at the bottom under their control through dependency for survival on the Cult. The whole structure is planned to be protected and enforced by a military-police state (Fig 4).

Here you have the reason for the global lockdowns of the fake pandemic to coldly destroy independent incomes and livelihoods and make everyone dependent on the 'state' (the Cult that controls the 'states'). I have warned in my books for many years about the plan to introduce a 'guaranteed income' – a barely survivable pittance – designed to impose dependency when employment was destroyed by AI technology and now even more comprehensively at great speed by the 'Covid' scam. Once the pandemic was played and lockdown consequences began to delete independent income the authorities began to talk right on cue about the need for a guaranteed income and a 'Great Reset'. Guaranteed income will be presented as benevolent governments seeking to help a desperate people – desperate as a direct result of actions of the same governments. The truth is that such payments are a trap. You will only get them if you do exactly what the authorities demand including mass vaccination (genetic manipulation). We have seen this theme already in Australia where those dependent on government benefits have them reduced if parents don't agree to have their children vaccinated according to an insane health-destroying government-dictated schedule. Calculated economic collapse applies to governments as well as people. The Cult wants rid of countries through the creation of a world state with countries broken up into regions ruled by a world government and super states like the European Union. Countries must be bankrupted, too, to this end and it's being achieved by the trillions in 'rescue packages' and furlough payments, trillions in lost taxation, and money-no-object spending on 'Covid' including constant all-medium advertising (programming) which has made the media dependent on government for much of its income. The day of reckoning is coming – as planned – for government spending and given that it has been made possible by printing money and not by production/taxation there is inflation on the way that has the

potential to wipe out monetary value. In that case there will be no need for the Cult to steal your money. It just won't be worth anything (see the German Weimar Republic before the Nazis took over). Many have been okay with lockdowns while getting a percentage of their income from so-called furlough payments without having to work. Those payments are dependent, however, on people having at least a theoretical job with a business considered non-essential and ordered to close. As these business go under because they are closed by lockdown after lockdown the furlough stops and it will for everyone eventually. Then what? The 'then what?' is precisely the idea.

Figure 4: The Hunger Games Society structure I have long warned was planned and now the 'Covid' hoax has made it possible. This is the real reason for lockdowns.

Hired hands

Between the Hunger Games Cult elite and the dependent population is planned to be a vicious military-police state (a fusion of the two into one force). This has been in the making for a long time with police looking ever more like the military and carrying weapons to match. The pandemic scam has seen this process accelerate so fast as

lockdown house arrest is brutally enforced by carefully recruited fascist minds and gormless system-servers. The police and military are planned to merge into a centrally-directed world army in a global structure headed by a world government which wouldn't be elected even by the election fixes now in place. The world army is not planned even to be human and instead wars would be fought, primarily against the population, using robot technology controlled by artificial intelligence. I have been warning about this for decades and now militaries around the world are being transformed by this very AI technology. The global regime that I describe is a particular form of fascism known as a technocracy in which decisions are not made by clueless and co-opted politicians but by unelected technocrats – scientists, engineers, technologists and bureaucrats. Cult-owned-and-controlled Silicon Valley giants are examples of technocracy and they already have far more power to direct world events than governments. They are with their censorship *selecting* governments. I know that some are calling the 'Great Reset' a Marxist communist takeover, but fascism and Marxism are different labels for the same tyranny. Tell those who lived in fascist Germany and Stalinist Russia that there was a difference in the way their freedom was deleted and their lives controlled. I could call it a fascist technocracy or a Marxist technocracy and they would be equally accurate. The Hunger Games society with its world government structure would oversee a world army, world central bank and single world cashless currency imposing its will on a microchipped population (Fig 5). Scan its different elements and see how the illusory pandemic is forcing society in this very direction at great speed. Leaders of 23 countries and the World Health Organization (WHO) backed the idea in March, 2021, of a global treaty for 'international cooperation' in 'health emergencies' and nations should 'come together as a global community for peaceful cooperation that extends beyond this crisis'. Cut the Orwellian bullshit and this means another step towards global government. The plan includes a cashless digital money system that I first warned about in 1993. Right at the start of 'Covid' the deeply corrupt Tedros

Adhanom Ghebreyesus, the crooked and merely gofer 'head' of the World Health Organization, said it was possible to catch the 'virus' by touching cash and it was better to use cashless means. The claim was ridiculous nonsense and like the whole 'Covid' mind-trick it was nothing to do with 'health' and everything to do with pushing every aspect of the Cult agenda. As a result of the Tedros lie the use of cash has plummeted. The Cult script involves a single world digital currency that would eventually be technologically embedded in the body. China is a massive global centre for the Cult and if you watch what is happening there you will know what is planned for everywhere. The Chinese government is developing a digital currency which would allow fines to be deducted immediately via AI for anyone caught on camera breaking its fantastic list of laws and the money is going to be programmable with an expiry date to ensure that no one can accrue wealth except the Cult and its operatives.

Figure 5: The structure of global control the Cult has been working towards for so long and this has been enormously advanced by the 'Covid' illusion.

Serfdom is so smart

The Cult plan is far wider, extreme, and more comprehensive than even most conspiracy researchers appreciate and I will come to the true depths of deceit and control in the chapters 'Who controls the

Cult?' and 'Escaping Wetiko'. Even the world that we know is crazy enough. We are being deluged with ever more sophisticated and controlling technology under the heading of 'smart'. We have smart televisions, smart meters, smart cards, smart cars, smart driving, smart roads, smart pills, smart patches, smart watches, smart skin, smart borders, smart pavements, smart streets, smart cities, smart communities, smart environments, smart growth, smart planet ... smart *everything* around us. Smart technologies and methods of operation are designed to interlock to create a global Smart Grid connecting the entirety of human society including human minds to create a centrally-dictated 'hive' mind. 'Smart cities' is code for densely-occupied megacities of total surveillance and control through AI. Ever more destructive frequency communication systems like 5G have been rolled out without any official testing for health and psychological effects (colossal). 5G/6G/7G systems are needed to run the Smart Grid and each one becomes more destructive of body and mind. Deleting independent income is crucial to forcing people into these AI-policed prisons by ending private property ownership (except for the Cult elite). The Cult's Great Reset now openly foresees a global society in which no one will own any possessions and everything will be rented while the Cult would own literally everything under the guise of government and corporations. The aim has been to use the lockdowns to destroy sources of income on a mass scale and when the people are destitute and in unrepayable amounts of debt (problem) Cult assets come forward with the pledge to write-off debt in return for handing over all property and possessions (solution). Everything – literally everything including people – would be connected to the Internet via AI. I was warning years ago about the coming Internet of Things (IoT) in which all devices and technology from your car to your fridge would be plugged into the Internet and controlled by AI. Now we are already there with much more to come. The next stage is the Internet of Everything (IoE) which is planned to include the connection of AI to the human brain and body to replace the human mind with a centrally-controlled AI mind. Instead of perceptions

being manipulated through control of information and censorship those perceptions would come direct from the Cult through AI. What do you think? You think whatever AI decides that you think. In human terms there would be no individual 'think' any longer. Too incredible? The ravings of a lunatic? Not at all. Cult-owned crazies in Silicon Valley have been telling us the plan for years without explaining the real motivation and calculated implications. These include Google executive and 'futurist' Ray Kurzweil who highlights the year 2030 for when this would be underway. He said:

Our thinking ... will be a hybrid of biological and non-biological thinking ... humans will be able to extend their limitations and 'think in the cloud' ... We're going to put gateways to the cloud in our brains ... We're going to gradually merge and enhance ourselves ... In my view, that's the nature of being human – we transcend our limitations.

As the technology becomes vastly superior to what we are then the small proportion that is still human gets smaller and smaller and smaller until it's just utterly negligible.

The sales-pitch of Kurzweil and Cult-owned Silicon Valley is that this would make us 'super-human' when the real aim is to make us post-human and no longer 'human' in the sense that we have come to know. The entire global population would be connected to AI and become the centrally-controlled 'hive-mind' of externally-delivered perceptions. The Smart Grid being installed to impose the Cult's will on the world is being constructed to allow particular locations – even one location – to control the whole global system. From these prime control centres, which absolutely include China and Israel, anything connected to the Internet would be switched on or off and manipulated at will. Energy systems could be cut, communication via the Internet taken down, computer-controlled driverless autonomous vehicles driven off the road, medical devices switched off, the potential is limitless given how much AI and Internet connections now run human society. We have seen nothing yet if we allow this to continue. Autonomous vehicle makers are working with law enforcement to produce cars designed to automatically pull over if they detect a police or emergency vehicle flashing from up to 100 feet away. At a police stop the car would be unlocked and the

window rolled down automatically. Vehicles would only take you where the computer (the state) allowed. The end of petrol vehicles and speed limiters on all new cars in the UK and EU from 2022 are steps leading to electric computerised transport over which ultimately you have no control. The picture is far bigger even than the Cult global network or web and that will become clear when I get to the nature of the 'spider'. There is a connection between all these happenings and the instigation of DNA-manipulating 'vaccines' (which aren't 'vaccines') justified by the 'Covid' hoax. That connection is the unfolding plan to transform the human body from a biological to a synthetic biological state and this is why synthetic biology is such a fast-emerging discipline of mainstream science. 'Covid vaccines' are infusing self-replicating synthetic genetic material into the cells to cumulatively take us on the Totalitarian Tiptoe from Human 1.0 to the synthetic biological Human 2.0 which will be physically and perceptually attached to the Smart Grid to one hundred percent control every thought, perception and deed. Humanity needs to wake up and *fast*.

This is the barest explanation of where the 'outcome' is planned to go but it's enough to see the journey happening all around us. Those new to this information will already see 'Covid' in a whole new context. I will add much more detail as we go along, but for the minutiae evidence see my mega-works, *The Answer, The Trigger* and *Everything You Need to Know But Have Never Been Told*.

Now – how does a Renegade Mind see the 'world'?

CHAPTER TWO

Renegade Perception

It is one thing to be clever and another to be wise
George R.R. Martin

A simple definition of the difference between a programmed mind and a Renegade Mind would be that one sees only dots while the other connects them to see the picture. Reading reality with accuracy requires the observer to (a) know the planned outcome and (b) realise that everything, but *everything*, is connected.

The entirety of infinite reality is connected – that's its very nature – and with human society an expression of infinite reality the same must apply. Simple cause and effect is a connection. The effect is triggered by the cause and the effect then becomes the cause of another effect. Nothing happens in isolation because it *can't*. Life in whatever reality is simple choice and consequence. We make choices and these lead to consequences. If we don't like the consequences we can make different choices and get different consequences which lead to other choices and consequences. The choice and the consequence are not only connected they are indivisible. You can't have one without the other as an old song goes. A few cannot control the world unless those being controlled allow that to happen – cause and effect, choice and consequence. Control – who has it and who doesn't – is a two-way process, a symbiotic relationship, involving the controller and controlled. 'They took my freedom away!!' Well, yes, but you also gave it to them. Humanity is

subjected to mass control because humanity has acquiesced to that control. This is all cause and effect and literally a case of give and take. In the same way world events of every kind are connected and the Cult works incessantly to sell the illusion of the random and coincidental to maintain the essential (to them) perception of dots that hide the picture. Renegade Minds know this and constantly scan the world for patterns of connection. This is absolutely pivotal in understanding the happenings in the world and without that perspective clarity is impossible. First you know the planned outcome and then you identify the steps on the journey – the day-by-day apparently random which, when connected in relation to the outcome, no longer appear as individual events, but as the proverbial *chain* of events leading in the same direction. I'll give you some examples:

Political puppet show

We are told to believe that politics is 'adversarial' in that different parties with different beliefs engage in an endless tussle for power. There may have been some truth in that up to a point – and only a point – but today divisions between 'different' parties are rhetorical not ideological. Even the rhetorical is fusing into one-speak as the parties eject any remaining free thinkers while others succumb to the ever-gathering intimidation of anyone with the 'wrong' opinion. The Cult is not a new phenomenon and can be traced back thousands of years as my books have documented. Its intergenerational initiates have been manipulating events with increasing effect the more that global power has been centralised. In ancient times the Cult secured control through the system of monarchy in which 'special' bloodlines (of which more later) demanded the right to rule as kings and queens simply by birthright and by vanquishing others who claimed the same birthright. There came a time, however, when people had matured enough to see the unfairness of such tyranny and demanded a say in who governed them. Note the word – *governed* them. Not served them – *governed* them, hence government defined as 'the political direction and control exercised over the

actions of the members, citizens, or inhabitants of communities, societies, and states; direction of the affairs of a state, community, etc.' Governments exercise control over rather than serve just like the monarchies before them. Bizarrely there are still countries like the United Kingdom which are ruled by a monarch *and* a government that officially answers to the monarch. The UK head of state and that of Commonwealth countries such as Canada, Australia and New Zealand is 'selected' by who in a *single family* had unprotected sex with whom and in what order. Pinch me it can't be true. Ouch! Shit, it is. The demise of monarchies in most countries offered a potential vacuum in which some form of free and fair society could arise and the Cult had that base covered. Monarchies had served its interests but they couldn't continue in the face of such widespread opposition and, anyway, replacing a 'royal' dictatorship that people could see with a dictatorship 'of the people' hiding behind the concept of 'democracy' presented far greater manipulative possibilities and ways of hiding coordinated tyranny behind the illusion of 'freedom'.

Democracy is quite wrongly defined as government selected by the population. This is not the case at all. It is government selected by *some* of the population (and then only in theory). This 'some' doesn't even have to be the majority as we have seen so often in first-past-the-post elections in which the so-called majority party wins fewer votes than the 'losing' parties combined. Democracy can give total power to a party in government from a minority of the votes cast. It's a sleight of hand to sell tyranny as freedom. Seventy-four million Trump-supporting Americans didn't vote for the 'Democratic' Party of Joe Biden in the distinctly dodgy election in 2020 and yet far from acknowledging the wishes and feelings of that great percentage of American society the Cult-owned Biden government set out from day one to destroy them and their right to a voice and opinion. Empty shell Biden and his Cult handlers said they were doing this to 'protect democracy'. Such is the level of lunacy and sickness to which politics has descended. Connect the dots and relate them to the desired outcome – a world government run by self-appointed technocrats and no longer even elected

politicians. While operating through its political agents in government the Cult is at the same time encouraging public distain for politicians by putting idiots and incompetents in theoretical power on the road to deleting them. The idea is to instil a public reaction that says of the technocrats: 'Well, they couldn't do any worse than the pathetic politicians.' It's all about controlling perception and Renegade Minds can see through that while programmed minds cannot when they are ignorant of both the planned outcome and the manipulation techniques employed to secure that end. This knowledge can be learned, however, and fast if people choose to get informed.

Politics may at first sight appear very difficult to control from a central point. I mean look at the 'different' parties and how would you be able to oversee them all and their constituent parts? In truth, it's very straightforward because of their structure. We are back to the pyramid of imposition and acquiescence. Organisations are structured in the same way as the system as a whole. Political parties are not open forums of free expression. They are hierarchies. I was a national spokesman for the British Green Party which claimed to be a different kind of politics in which influence and power was devolved; but I can tell you from direct experience – and it's far worse now – that Green parties are run as hierarchies like all the others however much they may try to hide that fact or kid themselves that it's not true. A very few at the top of all political parties are directing policy and personnel. They decide if you are elevated in the party or serve as a government minister and to do that you have to be a yes man or woman. Look at all the maverick political thinkers who never ascended the greasy pole. If you want to progress within the party or reach 'high-office' you need to fall into line and conform. Exceptions to this are rare indeed. Should you want to run for parliament or Congress you have to persuade the local or state level of the party to select you and for that you need to play the game as dictated by the hierarchy. If you secure election and wish to progress within the greater structure you need to go on conforming to what is acceptable to those running the hierarchy

from the peak of the pyramid. Political parties are perceptual gulags and the very fact that there are party 'Whips' appointed to 'whip' politicians into voting the way the hierarchy demands exposes the ridiculous idea that politicians are elected to serve the people they are supposed to represent. Cult operatives and manipulation has long seized control of major parties that have any chance of forming a government and at least most of those that haven't. A new party forms and the Cult goes to work to infiltrate and direct. This has reached such a level today that you see video compilations of 'leaders' of all parties whether Democrats, Republicans, Conservative, Labour and Green parroting the same Cult mantra of 'Build Back Better' and the 'Great Reset' which are straight off the Cult song-sheet to describe the transformation of global society in response to the Cult-instigated hoaxes of the 'Covid pandemic' and human-caused 'climate change'. To see Caroline Lucas, the Green Party MP that I knew when I was in the party in the 1980s, speaking in support of plans proposed by Cult operative Klaus Schwab representing the billionaire global elite is a real head-shaker.

Many parties – one master

The party system is another mind-trick and was instigated to change the nature of the dictatorship by swapping 'royalty' for dark suits that people believed – though now ever less so – represented their interests. Understanding this trick is to realise that a single force (the Cult) controls all parties either directly in terms of the major ones or through manipulation of perception and ideology with others. You don't need to manipulate Green parties to demand your transformation of society in the name of 'climate change' when they are obsessed with the lie that this is essential to 'save the planet'. You just give them a platform and away they go serving your interests while believing they are being environmentally virtuous. America's political structure is a perfect blueprint for how the two or multi-party system is really a one-party state. The Republican Party is controlled from one step back in the shadows by a group made up of billionaires and their gofers known as neoconservatives or Neocons.

I have exposed them in fine detail in my books and they were the driving force behind the policies of the imbecilic presidency of Boy George Bush which included 9/11 (see *The Trigger* for a comprehensive demolition of the official story), the subsequent 'war on terror' (war *of* terror) and the invasions of Afghanistan and Iraq. The latter was a No-Problem-Reaction-Solution based on claims by Cult operatives, including Bush and British Prime Minister Tony Blair, about Saddam Hussein's 'weapons of mass destruction' which did not exist as war criminals Bush and Blair well knew.

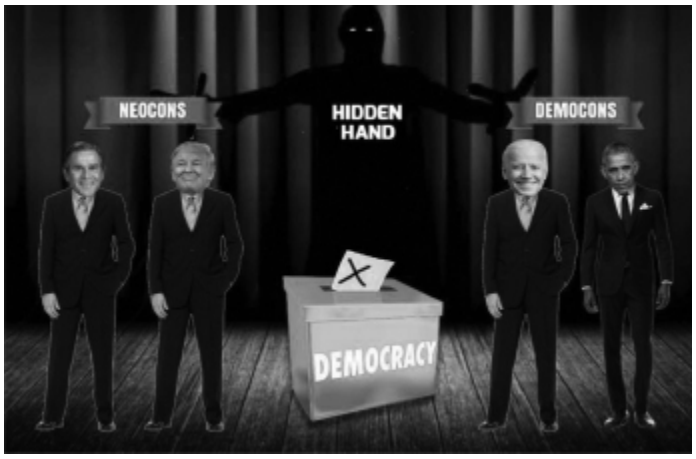


Figure 6: Different front people, different parties – same control system.

The Democratic Party has its own 'Neocon' group controlling from the background which I call the 'Democons' and here's the penny-drop – the Neocons and Democons answer to the same masters one step further back into the shadows (Fig 6). At that level of the Cult the Republican and Democrat parties are controlled by the same people and no matter which is in power the Cult is in power. This is how it works in almost every country and certainly in Britain with Conservative, Labour, Liberal Democrat and Green parties now all on the same page whatever the rhetoric may be in their feeble attempts to appear different. Neocons operated at the time of Bush through a think tank called The Project for the New American Century which in September, 2000, published a document entitled *Rebuilding America's Defenses: Strategies, Forces, and Resources*

For a New Century demanding that America fight 'multiple, simultaneous major theatre wars' as a 'core mission' to force regime-change in countries including Iraq, Libya and Syria. Neocons arranged for Bush ('Republican') and Blair ('Labour Party') to front-up the invasion of Iraq and when they departed the Democons orchestrated the targeting of Libya and Syria through Barack Obama ('Democrat') and British Prime Minister David Cameron ('Conservative Party'). We have 'different' parties and 'different' people, but the same unfolding script. The more the Cult has seized the reigns of parties and personnel the more their policies have transparently pursued the same agenda to the point where the fascist 'Covid' impositions of the Conservative junta of Jackboot Johnson in Britain were opposed by the Labour Party because they were not fascist enough. The Labour Party is likened to the US Democrats while the Conservative Party is akin to a British version of the Republicans and on both sides of the Atlantic they all speak the same language and support the direction demanded by the Cult although some more enthusiastically than others. It's a similar story in country after country because it's all centrally controlled. Oh, but what about Trump? I'll come to him shortly. Political 'choice' in the 'party' system goes like this: You vote for Party A and they get into government. You don't like what they do so next time you vote for Party B and they get into government. You don't like what they do when it's pretty much the same as Party A and why wouldn't that be with both controlled by the same force? Given that only two, sometimes three, parties have any chance of forming a government to get rid of Party B that you don't like you have to vote again for Party A which … you don't like. This, ladies and gentlemen, is what they call 'democracy' which we are told – wrongly – is a term interchangeable with 'freedom'.

The cult of cults

At this point I need to introduce a major expression of the Global Cult known as Sabbatian-Frankism. Sabbatian is also spelt as Sabbatean. I will summarise here. I have published major exposés

and detailed background in other works. Sabbatian-Frankism combines the names of two frauds posing as 'Jewish' men, Sabbatai Zevi (1626-1676), a rabbi, black magician and occultist who proclaimed he was the Jewish messiah; and Jacob Frank (1726-1791), the Polish 'Jew', black magician and occultist who said he was the reincarnation of 'messiah' Zevi and biblical patriarch Jacob. They worked across two centuries to establish the Sabbatian-Frankist cult that plays a major, indeed central, role in the manipulation of human society by the Global Cult which has its origins much further back in history than Sabbatai Zevi. I should emphasise two points here in response to the shrill voices that will scream 'anti-Semitism': (1) Sabbatian-Frankists are NOT Jewish and only pose as such to hide their cult behind a Jewish façade; and (2) my information about this cult has come from Jewish sources who have long realised that their society and community has been infiltrated and taken over by interloper Sabbatian-Frankists. Infiltration has been the foundation technique of Sabbatian-Frankism from its official origin in the 17th century. Zevi's Sabbatian sect attracted a massive following described as the biggest messianic movement in Jewish history, spreading as far as Africa and Asia, and he promised a return for the Jews to the 'Promised Land' of Israel. Sabbatianism was not Judaism but an inversion of everything that mainstream Judaism stood for. So much so that this sinister cult would have a feast day when Judaism had a fast day and whatever was forbidden in Judaism the Sabbatians were encouraged and even commanded to do. This included incest and what would be today called Satanism. Members were forbidden to marry outside the sect and there was a system of keeping their children ignorant of what they were part of until they were old enough to be trusted not to unknowingly reveal anything to outsiders. The same system is employed to this day by the Global Cult in general which Sabbatian-Frankism has enormously influenced and now largely controls.

Zevi and his Sabbatians suffered a setback with the intervention by the Sultan of the Islamic Ottoman Empire in the Middle East and what is now the Republic of Turkey where Zevi was located. The

Sultan gave him the choice of proving his 'divinity', converting to Islam or facing torture and death. Funnily enough Zevi chose to convert or at least appear to. Some of his supporters were disillusioned and drifted away, but many did not with 300 families also converting – only in theory – to Islam. They continued behind this Islamic smokescreen to follow the goals, rules and rituals of Sabbatianism and became known as 'crypto-Jews' or the 'Dönmeh' which means 'to turn'. This is rather ironic because they didn't 'turn' and instead hid behind a fake Islamic persona. The process of appearing to be one thing while being very much another would become the calling card of Sabbatianism especially after Zevi's death and the arrival of the Satanist Jacob Frank in the 18th century when the cult became Sabbatian-Frankism and plumbed still new depths of depravity and infiltration which included – still includes – human sacrifice and sex with children. Wherever Sabbatians go paedophilia and Satanism follow and is it really a surprise that Hollywood is so infested with child abuse and Satanism when it was established by Sabbatian-Frankists and is still controlled by them? Hollywood has been one of the prime vehicles for global perceptual programming and manipulation. How many believe the version of 'history' portrayed in movies when it is a travesty and inversion (again) of the truth? Rabbi Marvin Antelman describes Frankism in his book, *To Eliminate the Opiate*, as 'a movement of complete evil' while Jewish professor Gershom Scholem said of Frank in *The Messianic Idea in Judaism*: 'In all his actions [he was] a truly corrupt and degenerate individual ... one of the most frightening phenomena in the whole of Jewish history.' Frank was excommunicated by traditional rabbis, as was Zevi, but Frank was undeterred and enjoyed vital support from the House of Rothschild, the infamous banking dynasty whose inner-core are Sabbatian-Frankists and not Jews. Infiltration of the Roman Church and Vatican was instigated by Frank with many Dönmeh 'turning' again to convert to Roman Catholicism with a view to hijacking the reins of power. This was the ever-repeating modus operandi and continues to be so. Pose as an advocate of the religion, culture or country that you want to control and then

manipulate your people into the positions of authority and influence largely as advisers, administrators and Svengalis for those that appear to be in power. They did this with Judaism, Christianity (Christian Zionism is part of this), Islam and other religions and nations until Sabbatian-Frankism spanned the world as it does today.

Sabbatian Saudis and the terror network

One expression of the Sabbatian-Frankist Dönmeh within Islam is the ruling family of Saudi Arabia, the House of Saud, through which came the vile distortion of Islam known as Wahhabism. This is the violent creed followed by terrorist groups like Al-Qaeda and ISIS or Islamic State. Wahhabism is the hand-chopping, head-chopping 'religion' of Saudi Arabia which is used to keep the people in a constant state of fear so the interloper House of Saud can continue to rule. Al-Qaeda and Islamic State were lavishly funded by the House of Saud while being created and directed by the Sabbatian-Frankist network in the United States that operates through the Pentagon, CIA and the government in general of whichever 'party'. The front man for the establishment of Wahhabism in the middle of the 18th century was a Sabbatian-Frankist 'crypto-Jew' posing as Islamic called Muhammad ibn Abd al-Wahhab. His daughter would marry the son of Muhammad bin Saud who established the first Saudi state before his death in 1765 with support from the British Empire. Bin Saud's successors would establish modern Saudi Arabia in league with the British and Americans in 1932 which allowed them to seize control of Islam's major shrines in Mecca and Medina. They have dictated the direction of Sunni Islam ever since while Iran is the major centre of the Shiite version and here we have the source of at least the public conflict between them. The Sabbatian network has used its Wahhabi extremists to carry out Problem-Reaction-Solution terrorist attacks in the name of 'Al-Qaeda' and 'Islamic State' to justify a devastating 'war on terror', ever-increasing surveillance of the population and to terrify people into compliance. Another insight of the Renegade Mind is the streetwise understanding that

just because a country, location or people are attacked doesn't mean that those apparently representing that country, location or people are not behind the attackers. Often they are *orchestrating* the attacks because of the societal changes that can be then justified in the name of 'saving the population from terrorists'.

I show in great detail in *The Trigger* how Sabbatian-Frankists were the real perpetrators of 9/11 and not '19 Arab hijackers' who were blamed for what happened. Observe what was justified in the name of 9/11 alone in terms of Middle East invasions, mass surveillance and control that fulfilled the demands of the Project for the New American Century document published by the Sabbatian Neocons. What appear to be enemies are on the deep inside players on the same Sabbatian team. Israel and Arab 'royal' dictatorships are all ruled by Sabbatians and the recent peace agreements between Israel and Saudi Arabia, the United Arab Emirates (UAE) and others are only making formal what has always been the case behind the scenes. Palestinians who have been subjected to grotesque tyranny since Israel was bombed and terrorised into existence in 1948 have never stood a chance. Sabbatian-Frankists have controlled Israel (so the constant theme of violence and war which Sabbatians love) and they have controlled the Arab countries that Palestinians have looked to for real support that never comes. 'Royal families' of the Arab world in Saudi Arabia, Bahrain, UAE, etc., are all Sabbatians with allegiance to the aims of the cult and not what is best for their Arabic populations. They have stolen the oil and financial resources from their people by false claims to be 'royal dynasties' with a genetic right to rule and by employing vicious militaries to impose their will.

Satanic 'illumination'

The Satanist Jacob Frank formed an alliance in 1773 with two other Sabbatians, Mayer Amschel Rothschild (1744-1812), founder of the Rothschild banking dynasty, and Jesuit-educated fraudulent Jew, Adam Weishaupt, and this led to the formation of the Bavarian Illuminati, firstly under another name, in 1776. The Illuminati would

be the manipulating force behind the French Revolution (1789-1799) and was also involved in the American Revolution (1775-1783) before and after the Illuminati's official creation. Weishaupt would later become (in public) a Protestant Christian in archetypal Sabbatian style. I read that his name can be decoded as Adam-Weishaupt or 'the first man to lead those who know'. He wasn't a leader in the sense that he was a subordinate, but he did lead those below him in a crusade of transforming human society that still continues today. The theme was confirmed as early as 1785 when a horseman courier called Lanz was reported to be struck by lighting and extensive Illuminati documents were found in his saddlebags. They made the link to Weishaupt and detailed the plan for world takeover. Current events with 'Covid' fascism have been in the making for a very long time. Jacob Frank was jailed for 13 years by the Catholic Inquisition after his arrest in 1760 and on his release he headed for Frankfurt, Germany, home city and headquarters of the House of Rothschild where the alliance was struck with Mayer Amschel Rothschild and Weishaupt. Rothschild arranged for Frank to be given the title of Baron and he became a wealthy nobleman with a big following of Jews in Germany, the Austro-Hungarian Empire and other European countries. Most of them would have believed he was on their side.

The name 'Illuminati' came from the Zohar which is a body of works in the Jewish mystical 'bible' called the Kabbalah. 'Zohar' is the foundation of Sabbatian-Frankist belief and in Hebrew 'Zohar' means 'splendour', 'radiance', 'illuminated', and so we have 'Illuminati'. They claim to be the 'Illuminated Ones' from their knowledge systematically hidden from the human population and passed on through generations of carefully-chosen initiates in the global secret society network or Cult. Hidden knowledge includes an awareness of the Cult agenda for the world and the nature of our collective reality that I will explore later. Cult 'illumination' is symbolised by the torch held by the Statue of Liberty which was gifted to New York by French Freemasons in Paris who knew exactly what it represents. 'Liberty' symbolises the goddess worshipped in

Babylon as Queen Semiramis or Ishtar. The significance of this will become clear. Notice again the ubiquitous theme of inversion with the Statue of 'Liberty' really symbolising mass control (Fig 7). A mirror-image statute stands on an island in the River Seine in Paris from where New York Liberty originated (Fig 8). A large replica of the Liberty flame stands on top of the Pont de l'Alma tunnel in Paris where Princess Diana died in a Cult ritual described in *The Biggest Secret*. Lucifer 'the light bringer' is related to all this (and much more as we'll see) and 'Lucifer' is a central figure in Sabbatian-Frankism and its associated Satanism. Sabbatians reject the Jewish Torah, or Pentateuch, the 'five books of Moses' in the Old Testament known as Genesis, Exodus, Leviticus, Numbers, and Deuteronomy which are claimed by Judaism and Christianity to have been dictated by 'God' to Moses on Mount Sinai. Sabbatians say these do not apply to them and they seek to replace them with the Zohar to absorb Judaism and its followers into their inversion which is an expression of a much greater global inversion. They want to delete all religions and force humanity to worship a one-world religion – Sabbatian Satanism that also includes worship of the Earth goddess. Satanic themes are being more and more introduced into mainstream society and while Christianity is currently the foremost target for destruction the others are planned to follow.

Figure 7: The Cult goddess of Babylon disguised as the Statue of Liberty holding the flame of Lucifer the 'light bringer'.

Figure 8: Liberty's mirror image in Paris where the New York version originated.

Marx brothers

Rabbi Marvin Antelman connects the Illuminati to the Jacobins in *To Eliminate the Opiate* and Jacobins were the force behind the French Revolution. He links both to the Bund der Gerechten, or League of the Just, which was the network that inflicted communism/Marxism on the world. Antelman wrote:

> The original inner circle of the Bund der Gerechten consisted of born Catholics, Protestants and Jews [Sabbatian-Frankist infiltrators], and those representatives of respective subdivisions formulated schemes for the ultimate destruction of their faiths. The heretical Catholics laid plans which they felt would take a century or more for the ultimate destruction of the church; the apostate Jews for the ultimate destruction of the Jewish religion.

Sabbatian-created communism connects into this anti-religion agenda in that communism does not allow for the free practice of religion. The Sabbatian 'Bund' became the International Communist Party and Communist League and in 1848 'Marxism' was born with the Communist Manifesto of Sabbatian assets Karl Marx and Friedrich Engels. It is absolutely no coincidence that Marxism, just a different name for fascist and other centrally-controlled tyrannies, is being imposed worldwide as a result of the 'Covid' hoax and nor that Marxist/fascist China was the place where the hoax originated. The reason for this will become very clear in the chapter 'Covid: The calculated catastrophe'. The so-called 'Woke' mentality has hijacked

traditional beliefs of the political left and replaced them with far-right make-believe 'social justice' better known as Marxism. Woke will, however, be swallowed by its own perceived 'revolution' which is really the work of billionaires and billionaire corporations feigning being 'Woke'. Marxism is being touted by Wokers as a replacement for 'capitalism' when we don't have 'capitalism'. We have cartelism in which the market is stitched up by the very Cult billionaires and corporations bankrolling Woke. Billionaires love Marxism which keeps the people in servitude while they control from the top. Terminally naïve Wokers think they are 'changing the world' when it's the Cult that is doing the changing and when they have played their vital part and become surplus to requirements they, too, will be targeted. The Illuminati-Jacobins were behind the period known as 'The Terror' in the French Revolution in 1793 and 1794 when Jacobin Maximillian de Robespierre and his Orwellian 'Committee of Public Safety' killed 17,000 'enemies of the Revolution' who had once been 'friends of the Revolution'. Karl Marx (1818-1883), whose Sabbatian creed of Marxism has cost the lives of at least 100 million people, is a hero once again to Wokers who have been systematically kept ignorant of real history by their 'education' programming. As a result they now promote a Sabbatian 'Marxist' abomination destined at some point to consume them. Rabbi Antelman, who spent decades researching the Sabbatian plot, said of the League of the Just and Karl Marx:

> Contrary to popular opinion Karl Marx did not originate the Communist Manifesto. He was paid for his services by the League of the Just, which was known in its country of origin, Germany, as the Bund der Geaechteten.

Antelman said the text attributed to Marx was the work of other people and Marx 'was only repeating what others already said'. Marx was 'a hired hack – lackey of the wealthy Illuminists'. Marx famously said that religion was the 'opium of the people' (part of the Sabbatian plan to demonise religion) and Antelman called his books, *To Eliminate the Opiate*. Marx was born Jewish, but his family converted to Christianity (Sabbatian modus operandi) and he

attacked Jews, not least in his book, *A World Without Jews*. In doing so he supported the Sabbatian plan to destroy traditional Jewishness and Judaism which we are clearly seeing today with the vindictive targeting of orthodox Jews by the Sabbatian government of Israel over 'Covid' laws. I don't follow any religion and it has done much damage to the world over centuries and acted as a perceptual straightjacket. Renegade Minds, however, are always asking *why* something is being done. It doesn't matter if they agree or disagree with what is happening – *why* is it happening is the question. The 'why?' can be answered with regard to religion in that religions create interacting communities of believers when the Cult wants to dismantle all discourse, unity and interaction (see 'Covid' lockdowns) and the ultimate goal is to delete all religions for a one-world religion of Cult Satanism worshipping their 'god' of which more later. We see the same 'why?' with gun control in America. I don't have guns and don't want them, but why is the Cult seeking to disarm the population at the same time that law enforcement agencies are armed to their molars and why has every tyrant in history sought to disarm people before launching the final takeover? They include Hitler, Stalin, Pol Pot and Mao who followed confiscation with violent seizing of power. You know it's a Cult agenda by the people who immediately race to the microphones to exploit dead people in multiple shootings. Ultra-Zionist Cult lackey Senator Chuck Schumer was straight on the case after ten people were killed in Boulder, Colorado in March, 2121. Simple rule … if Schumer wants it the Cult wants it and the same with his ultra-Zionist mate the wild-eyed Senator Adam Schiff. At the same time they were calling for the disarmament of Americans, many of whom live a long way from a police response, Schumer, Schiff and the rest of these pampered clowns were sitting on Capitol Hill behind a razor-wired security fence protected by thousands of armed troops in addition to their own armed bodyguards. Mom and pop in an isolated home? They're just potential mass shooters.

Zion Mainframe

Sabbatian-Frankists and most importantly the Rothschilds were behind the creation of 'Zionism', a political movement that demanded a Jewish homeland in Israel as promised by Sabbatai Zevi. The very symbol of Israel comes from the German meaning of the name Rothschild. Dynasty founder Mayer Amschel Rothschild changed the family name from Bauer to Rothschild, or 'Red-Shield' in German, in deference to the six-pointed 'Star of David' hexagram displayed on the family's home in Frankfurt. The symbol later appeared on the flag of Israel after the Rothschilds were centrally involved in its creation. Hexagrams are not a uniquely Jewish symbol and are widely used in occult ('hidden') networks often as a symbol for Saturn (see my other books for why). Neither are Zionism and Jewishness interchangeable. Zionism is a political movement and philosophy and not a 'race' or a people. Many Jews oppose Zionism and many non-Jews, including US President Joe Biden, call themselves Zionists as does Israel-centric Donald Trump. America's support for the Israel government is pretty much a gimme with ultra-Zionist billionaires and corporations providing fantastic and dominant funding for both political parties. Former Congresswoman Cynthia McKinney has told how she was approached immediately she ran for office to 'sign the pledge' to Israel and confirm that she would always vote in that country's best interests. All American politicians are approached in this way. Anyone who refuses will get no support or funding from the enormous and all-powerful Zionist lobby that includes organisations like mega-lobby group AIPAC, the American Israel Public Affairs Committee. Trump's biggest funder was ultra-Zionist casino and media billionaire Sheldon Adelson while major funders of the Democratic Party include ultra-Zionist George Soros and ultra-Zionist financial and media mogul, Haim Saban. Some may reel back at the suggestion that Soros is an Israel-firster (Sabbatian-controlled Israel-firster), but Renegade Minds watch the actions not the words and everywhere Soros donates his billions the Sabbatian agenda benefits. In the spirit of Sabbatian inversion Soros pledged $1 billion for a new university network to promote 'liberal values and tackle intolerance'. He made the announcement during his annual speech

at the Cult-owned World Economic Forum in Davos, Switzerland, in January, 2020, after his 'harsh criticism' of 'authoritarian rulers' around the world. You can only laugh at such brazen mendacity. How *he* doesn't laugh is the mystery. Translated from the Orwellian 'liberal values and tackle intolerance' means teaching non-white people to hate white people and for white people to loathe themselves for being born white. The reason for that will become clear.

The 'Anti-Semitism' fraud

Zionists support the Jewish homeland in the land of Palestine which has been the Sabbatian-Rothschild goal for so long, but not for the benefit of Jews. Sabbatians and their global Anti-Semitism Industry have skewed public and political opinion to equate opposing the violent extremes of Zionism to be a blanket attack and condemnation of all Jewish people. Sabbatians and their global Anti-Semitism Industry have skewed public and political opinion to equate opposing the violent extremes of Zionism to be a blanket attack and condemnation of all Jewish people. This is nothing more than a Sabbatian protection racket to stop legitimate investigation and exposure of their agendas and activities. The official definition of 'anti-Semitism' has more recently been expanded to include criticism of Zionism – a *political movement* – and this was done to further stop exposure of Sabbatian infiltrators who created Zionism as we know it today in the 19th century. Renegade Minds will talk about these subjects when they know the shit that will come their way. People must decide if they want to know the truth or just cower in the corner in fear of what others will say. Sabbatians have been trying to label me as 'anti-Semitic' since the 1990s as I have uncovered more and more about their background and agendas. Useless, gutless, fraudulent 'journalists' then just repeat the smears without question and on the day I was writing this section a pair of unquestioning repeaters called Ben Quinn and Archie Bland (how appropriate) outright called me an 'anti-Semite' in the establishment propaganda sheet, the London *Guardian*, with no supporting evidence. The

Sabbatian Anti-Semitism Industry said so and who are they to question that? They wouldn't dare. Ironically 'Semitic' refers to a group of languages in the Middle East that are almost entirely Arabic. 'Anti-Semitism' becomes 'anti-Arab' which if the consequences of this misunderstanding were not so grave would be hilarious. Don't bother telling Quinn and Bland. I don't want to confuse them, bless 'em. One reason I am dubbed 'anti-Semitic' is that I wrote in the 1990s that Jewish operatives (Sabbatians) were heavily involved in the Russian Revolution when Sabbatians overthrew the Romanov dynasty. This apparently made me 'anti-Semitic'. Oh, really? Here is a section from *The Trigger*:

British journalist Robert Wilton confirmed these themes in his 1920 book *The Last Days of the Romanovs* when he studied official documents from the Russian government to identify the members of the Bolshevik ruling elite between 1917 and 1919. The Central Committee included 41 Jews among 62 members; the Council of the People's Commissars had 17 Jews out of 22 members; and 458 of the 556 most important Bolshevik positions between 1918 and 1919 were occupied by Jewish people. Only 17 were Russian. Then there were the 23 Jews among the 36 members of the vicious Cheka Soviet secret police established in 1917 who would soon appear all across the country.

Professor Robert Service of Oxford University, an expert on 20th century Russian history, found evidence that ['Jewish'] Leon Trotsky had sought to make sure that Jews were enrolled in the Red Army and were disproportionately represented in the Soviet civil bureaucracy that included the Cheka which performed mass arrests, imprisonment and executions of 'enemies of the people'. A US State Department Decimal File (861.00/5339) dated November 13th, 1918, names [Rothschild banking agent in America] Jacob Schiff and a list of ultra-Zionists as funders of the Russian Revolution leading to claims of a 'Jewish plot', but the key point missed by all is they were not 'Jews' – they were Sabbatian-Frankists.

Britain's Winston Churchill made the same error by mistake or otherwise. He wrote in a 1920 edition of the *Illustrated Sunday Herald* that those behind the Russian revolution were part of a 'worldwide conspiracy for the overthrow of civilisation and for the reconstitution of society on the basis of arrested development, of envious malevolence, and impossible equality' (see 'Woke' today because that has been created by the same network). Churchill said there was no need to exaggerate the part played in the creation of Bolshevism and in the actual bringing about of the Russian

Revolution 'by these international and for the most part atheistical Jews' ['atheistical Jews' = Sabbatians]. Churchill said it is certainly a very great one and probably outweighs all others: 'With the notable exception of Lenin, the majority of the leading figures are Jews.' He went on to describe, knowingly or not, the Sabbatian modus operandi of placing puppet leaders nominally in power while they control from the background:

> Moreover, the principal inspiration and driving power comes from the Jewish leaders. Thus Tchitcherin, a pure Russian, is eclipsed by his nominal subordinate, Litvinoff, and the influence of Russians like Bukharin or Lunacharski cannot be compared with the power of Trotsky, or of Zinovieff, the Dictator of the Red Citadel (Petrograd), or of Krassin or Radek – all Jews. In the Soviet institutions the predominance of Jews is even more astonishing. And the prominent, if not indeed the principal, part in the system of terrorism applied by the Extraordinary Commissions for Combatting Counter-Revolution has been taken by Jews, and in some notable cases by Jewesses.

What I said about seriously disproportionate involvement in the Russian Revolution by Jewish 'revolutionaries' (Sabbatians) is provable fact, but truth is no defence against the Sabbatian Anti-Semitism Industry, its repeater parrots like Quinn and Bland, and the now breathtaking network of so-called 'Woke' 'anti-hate' groups with interlocking leaderships and funding which have the role of discrediting and silencing anyone who gets too close to exposing the Sabbatians. We have seen 'truth is no defence' confirmed in legal judgements with the Saskatchewan Human Rights Commission in Canada decreeing this: 'Truthful statements can be presented in a manner that would meet the definition of hate speech, and not all truthful statements must be free from restriction.' Most 'anti-hate' activists, who are themselves consumed by hatred, are too stupid and ignorant of the world to know how they are being used. They are far too far up their own virtue-signalling arses and it's far too dark for them to see anything.

The 'revolution' game

The background and methods of the 'Russian' Revolution are straight from the Sabbatian playbook seen in the French Revolution

and endless others around the world that appear to start as a revolution of the people against tyrannical rule and end up with a regime change to more tyrannical rule overtly or covertly. Wars, terror attacks and regime overthrows follow the Sabbatian cult through history with its agents creating them as Problem-Reaction-Solutions to remove opposition on the road to world domination. Sabbatian dots connect the Rothschilds with the Illuminati, Jacobins of the French Revolution, the 'Bund' or League of the Just, the International Communist Party, Communist League and the Communist Manifesto of Karl Marx and Friedrich Engels that would lead to the Rothschild-funded Russian Revolution. The sequence comes under the heading of 'creative destruction' when you advance to your global goal by continually destroying the status quo to install a new status quo which you then also destroy. The two world wars come to mind. With each new status quo you move closer to your planned outcome. Wars and mass murder are to Sabbatians a collective blood sacrifice ritual. They are obsessed with death for many reasons and one is that death is an inversion of life. Satanists and Sabbatians are obsessed with death and often target churches and churchyards for their rituals. Inversion-obsessed Sabbatians explain the use of inverted symbolism including the *inverted* pentagram and *inverted* cross. The inversion of the cross has been related to targeting Christianity, but the cross was a religious symbol long before Christianity and its inversion is a statement about the Sabbatian mentality and goals more than any single religion.

Sabbatians operating in Germany were behind the rise of the occult-obsessed Nazis and the subsequent Jewish exodus from Germany and Europe to Palestine and the United States after World War Two. The Rothschild dynasty was at the forefront of this both as political manipulators and by funding the operation. Why would Sabbatians help to orchestrate the horrors inflicted on Jews by the Nazis and by Stalin after they organised the Russian Revolution? Sabbatians hate Jews and their religion, that's why. They pose as Jews and secure positions of control within Jewish society and play the 'anti-Semitism' card to protect themselves from exposure

through a global network of organisations answering to the Sabbatian-created-and-controlled globe-spanning intelligence network that involves a stunning web of military-intelligence operatives and operations for a tiny country of just nine million. Among them are Jewish assets who are not Sabbatians but have been convinced by them that what they are doing is for the good of Israel and the Jewish community to protect them from what they have been programmed since childhood to believe is a Jew-hating hostile world. The Jewish community is just a highly convenient cover to hide the true nature of Sabbatians. Anyone getting close to exposing their game is accused by Sabbatian place-people and gofers of 'anti-Semitism' and claiming that all Jews are part of a plot to take over the world. I am not saying that. I am saying that Sabbatians – the *real* Jew-haters – have infiltrated the Jewish community to use them both as a cover and an 'anti-Semitic' defence against exposure. Thus we have the Anti-Semitism Industry targeted researchers in this way and most Jewish people think this is justified and genuine. They don't know that their 'Jewish' leaders and institutions of state, intelligence and military are not controlled by Jews at all, but cultists and stooges of Sabbatian-Frankism. I once added my name to a pro-Jewish freedom petition online and the next time I looked my name was gone and text had been added to the petition blurb to attack me as an 'anti-Semite' such is the scale of perceptual programming.

Moving on America

I tell the story in *The Trigger* and a chapter called 'Atlantic Crossing' how particularly after Israel was established the Sabbatians moved in on the United States and eventually grasped control of government administration, the political system via both Democrats and Republicans, the intelligence community like the CIA and National Security Agency (NSA), the Pentagon and mass media. Through this seriously compartmentalised network Sabbatians and their operatives in Mossad, Israeli Defense Forces (IDF) and US agencies pulled off 9/11 and blamed it on 19 'Al-Qaeda hijackers' dominated by men from, or connected to, Sabbatian-ruled Saudi

Arabia. The '19' were not even on the planes let alone flew those big passenger jets into buildings while being largely incompetent at piloting one-engine light aircraft. 'Hijacker' Hani Hanjour who is said to have flown American Airlines Flight 77 into the Pentagon with a turn and manoeuvre most professional pilots said they would have struggled to do was banned from renting a small plane by instructors at the Freeway Airport in Bowie, Maryland, just *six weeks* earlier on the grounds that he was an incompetent pilot. The Jewish population of the world is just 0.2 percent with even that almost entirely concentrated in Israel (75 percent Jewish) and the United States (around two percent). This two percent and globally 0.2 percent refers to *Jewish* people and not Sabbatian interlopers who are a fraction of that fraction. What a sobering thought when you think of the fantastic influence on world affairs of tiny Israel and that the Project for the New America Century (PNAC) which laid out the blueprint in September, 2000, for America's war on terror and regime change wars in Iraq, Libya and Syria was founded and dominated by Sabbatians known as 'Neocons'. The document conceded that this plan would not be supported politically or publicly without a major attack on American soil and a Problem-Reaction-Solution excuse to send troops to war across the Middle East. Sabbatian Neocons said:

> ... [The] process of transformation ... [war and regime change] ... is likely to be a long one, absent some catastrophic and catalysing event – like a new Pearl Harbor.

Four months later many of those who produced that document came to power with their inane puppet George Bush from the long-time Sabbatian Bush family. They included Sabbatian Dick Cheney who was officially vice-president, but really de-facto president for the entirety of the 'Bush' government. Nine months after the 'Bush' inauguration came what Bush called at the time 'the Pearl Harbor of the 21st century' and with typical Sabbatian timing and symbolism 2001 was the 60th anniversary of the attack in 1941 by the Japanese Air Force on Pearl Harbor, Hawaii, which allowed President Franklin Delano Roosevelt to take the United States into a Sabbatian-

instigated Second World War that he said in his election campaign that he never would. The evidence is overwhelming that Roosevelt and his military and intelligence networks knew the attack was coming and did nothing to stop it, but they did make sure that America's most essential naval ships were not in Hawaii at the time. Three thousand Americans died in the Pearl Harbor attacks as they did on September 11th. By the 9/11 year of 2001 Sabbatians had widely infiltrated the US government, military and intelligence operations and used their compartmentalised assets to pull off the 'Al-Qaeda' attacks. If you read *The Trigger* it will blow your mind to see the utterly staggering concentration of 'Jewish' operatives (Sabbatian infiltrators) in essential positions of political, security, legal, law enforcement, financial and business power before, during, and after the attacks to make them happen, carry them out, and then cover their tracks – and I do mean *staggering* when you think of that 0.2 percent of the world population and two percent of Americans which are Jewish while Sabbatian infiltrators are a fraction of that. A central foundation of the 9/11 conspiracy was the hijacking of government, military, Air Force and intelligence computer systems in real time through 'back-door' access made possible by Israeli (Sabbatian) 'cyber security' software. Sabbatian-controlled Israel is on the way to rivalling Silicon Valley for domination of cyberspace and is becoming the dominant force in cyber-security which gives them access to entire computer systems and their passcodes across the world. Then add to this that Zionists head (officially) Silicon Valley giants like Google (Larry Page and Sergey Brin), Google-owned YouTube (Susan Wojcicki), Facebook (Mark Zuckerberg and Sheryl Sandberg), and Apple (Chairman Arthur D. Levinson), and that ultra-Zionist hedge fund billionaire Paul Singer has a $1 billion stake in Twitter which is only nominally headed by 'CEO' pothead Jack Dorsey. As cable news host Tucker Carlson said of Dorsey: 'There used to be debate in the medical community whether dropping a ton of acid had permanent effects and I think that debate has now ended.' Carlson made the comment after Dorsey told a hearing on Capitol Hill (if you cut through his bullshit) that he

believed in free speech so long as he got to decide what you can hear and see. These 'big names' of Silicon Valley are only front men and women for the Global Cult, not least the Sabbatians, who are the true controllers of these corporations. Does anyone still wonder why these same people and companies have been ferociously censoring and banning people (like me) for exposing any aspect of the Cult agenda and especially the truth about the 'Covid' hoax which Sabbatians have orchestrated?

The Jeffrey Epstein paedophile ring was a Sabbatian operation. He was officially 'Jewish' but he was a Sabbatian and women abused by the ring have told me about the high number of 'Jewish' people involved. The Epstein horror has Sabbatian written all over it and matches perfectly their modus operandi and obsession with sex and ritual. Epstein was running a Sabbatian blackmail ring in which famous people with political and other influence were provided with young girls for sex while everything was being filmed and recorded on hidden cameras and microphones at his New York house, Caribbean island and other properties. Epstein survivors have described this surveillance system to me and some have gone public. Once the famous politician or other figure knew he or she was on video they tended to do whatever they were told. Here we go again …when you've got them by the balls their hearts and minds will follow. Sabbatians use this blackmail technique on a wide scale across the world to entrap politicians and others they need to act as demanded. Epstein's private plane, the infamous 'Lolita Express', had many well-known passengers including Bill Clinton while Bill Gates has flown on an Epstein plane and met with him four years after Epstein had been jailed for paedophilia. They subsequently met many times at Epstein's home in New York according to a witness who was there. Epstein's infamous side-kick was Ghislaine Maxwell, daughter of Mossad agent and ultra-Zionist mega-crooked British businessman, Bob Maxwell, who at one time owned the *Daily Mirror* newspaper. Maxwell was murdered at sea on his boat in 1991 by Sabbatian-controlled Mossad when he became a liability with his

business empire collapsing as a former Mossad operative has confirmed (see *The Trigger*).

Money, money, money, funny money …

Before I come to the Sabbatian connection with the last three US presidents I will lay out the crucial importance to Sabbatians of controlling banking and finance. Sabbatian Mayer Amschel Rothschild set out to dominate this arena in his family's quest for total global control. What is freedom? It is, in effect, choice. The more choices you have the freer you are and the fewer your choices the more you are enslaved. In the global structure created over centuries by Sabbatians the biggest decider and restrictor of choice is … money. Across the world if you ask people what they would like to do with their lives and why they are not doing that they will reply 'I don't have the money'. This is the idea. A global elite of multi-billionaires are described as 'greedy' and that is true on one level; but control of money – who has it and who doesn't – is not primarily about greed. It's about control. Sabbatians have seized ever more control of finance and sucked the wealth of the world out of the hands of the population. We talk now, after all, about the 'One-percent' and even then the wealthiest are a lot fewer even than that. This has been made possible by a money scam so outrageous and so vast it could rightly be called the scam of scams founded on creating 'money' out of nothing and 'loaning' that with interest to the population. Money out of nothing is called 'credit'. Sabbatians have asserted control over governments and banking ever more completely through the centuries and secured financial laws that allow banks to lend hugely more than they have on deposit in a confidence trick known as fractional reserve lending. Imagine if you could lend money that doesn't exist and charge the recipient interest for doing so. You would end up in jail. Bankers by contrast end up in mansions, private jets, Malibu and Monaco.

Banks are only required to keep a fraction of their deposits and wealth in their vaults and they are allowed to lend 'money' they don't have called 'credit. Go into a bank for a loan and if you succeed

the banker will not move any real wealth into your account. They will type into your account the amount of the agreed 'loan' – say £100,000. This is not wealth that really exists; it is non-existent, fresh-air, created-out-of-nothing 'credit' which has never, does not, and will never exist except in theory. Credit is backed by nothing except wind and only has buying power because people think that it has buying power and accept it in return for property, goods and services. I have described this situation as like those cartoon characters you see chasing each other and when they run over the edge of a cliff they keep running forward on fresh air until one of them looks down, realises what's happened, and they all crash into the ravine. The whole foundation of the Sabbatian financial system is to stop people looking down except for periodic moments when they want to crash the system (as in 2008 and 2020 ongoing) and reap the rewards from all the property, businesses and wealth their borrowers had signed over as 'collateral' in return for a 'loan' of fresh air. Most people think that money is somehow created by governments when it comes into existence from the start as a debt through banks 'lending' illusory money called credit. Yes, the very currency of exchange is a *debt* from day one issued as an interest-bearing loan. Why don't governments create money interest-free and lend it to their people interest-free? Governments are controlled by Sabbatians and the financial system is controlled by Sabbatians for whom interest-free money would be a nightmare come true. Sabbatians underpin their financial domination through their global network of central banks, including the privately-owned US Federal Reserve and Britain's Bank of England, and this is orchestrated by a privately-owned central bank coordination body called the Bank for International Settlements in Basle, Switzerland, created by the usual suspects including the Rockefellers and Rothschilds. Central bank chiefs don't answer to governments or the people. They answer to the Bank for International Settlements or, in other words, the Global Cult which is dominated today by Sabbatians.

Built-in disaster

There are so many constituent scams within the overall banking scam. When you take out a loan of thin-air credit only the amount of that loan is theoretically brought into circulation to add to the amount in circulation; but you are paying back the principle plus interest. The additional interest is not created and this means that with every 'loan' there is a shortfall in the money in circulation between what is borrowed and what has to be paid back. There is never even close to enough money in circulation to repay all outstanding public and private debt including interest. Coldly weaved in the very fabric of the system is the certainty that some will lose their homes, businesses and possessions to the banking 'lender'. This is less obvious in times of 'boom' when the amount of money in circulation (and the debt) is expanding through more people wanting and getting loans. When a downturn comes and the money supply contracts it becomes painfully obvious that there is not enough money to service all debt and interest. This is less obvious in times of 'boom' when the amount of money in circulation (and the debt) is expanding through more people wanting and getting loans. When a downturn comes and the money supply contracts and it becomes painfully obvious – as in 2008 and currently – that there is not enough money to service all debt and interest. Sabbatian banksters have been leading the human population through a calculated series of booms (more debt incurred) and busts (when the debt can't be repaid and the banks get the debtor's tangible wealth in exchange for non-existent 'credit'). With each 'bust' Sabbatian bankers have absorbed more of the world's tangible wealth and we end up with the One-percent. Governments are in bankruptcy levels of debt to the same system and are therefore owned by a system they do not control. The Federal Reserve, 'America's central bank', is privately-owned and American presidents only nominally appoint its chairman or woman to maintain the illusion that it's an arm of government. It's not. The 'Fed' is a cartel of private banks which handed billions to its associates and friends after the crash of 2008 and has been Sabbatian-controlled since it was manipulated into being in 1913 through the covert trickery of Rothschild banking agents Jacob Schiff and Paul

Warburg, and the Sabbatian Rockefeller family. Somehow from a Jewish population of two-percent and globally 0.2 percent (Sabbatian interlopers remember are far smaller) ultra-Zionists headed the Federal Reserve for 31 years between 1987 and 2018 in the form of Alan Greenspan, Bernard Bernanke and Janet Yellen (now Biden's Treasury Secretary) with Yellen's deputy chairman a Israeli-American duel citizen and ultra-Zionist Stanley Fischer, a former governor of the Bank of Israel. Ultra-Zionist Fed chiefs spanned the presidencies of Ronald Reagan ('Republican'), Father George Bush ('Republican'), Bill Clinton ('Democrat'), Boy George Bush ('Republican') and Barack Obama ('Democrat'). We should really add the pre-Greenspan chairman, Paul Adolph Volcker, 'appointed' by Jimmy Carter ('Democrat') who ran the Fed between 1979 and 1987 during the Carter and Reagan administrations before Greenspan took over. Volcker was a long-time associate and business partner of the Rothschilds. No matter what the 'party' officially in power the United States economy was directed by the same force. Here are members of the Obama, Trump and Biden administrations and see if you can make out a common theme.

Barack Obama ('Democrat')

Ultra-Zionists Robert Rubin, Larry Summers, and Timothy Geithner ran the US Treasury in the Clinton administration and two of them reappeared with Obama. Ultra-Zionist Fed chairman Alan Greenspan had manipulated the crash of 2008 through deregulation and jumped ship just before the disaster to make way for ultra-Zionist Bernard Bernanke to hand out trillions to Sabbatian 'too big to fail' banks and businesses, including the ubiquitous ultra-Zionist Goldman Sachs which has an ongoing staff revolving door operation between itself and major financial positions in government worldwide. Obama inherited the fallout of the crash when he took office in January, 2009, and fortunately he had the support of his ultra-Zionist White House Chief of Staff Rahm Emmanuel, son of a terrorist who helped to bomb Israel into being in 1948, and his ultra-Zionist senior adviser David Axelrod, chief strategist in Obama's two

successful presidential campaigns. Emmanuel, later mayor of Chicago and former senior fundraiser and strategist for Bill Clinton, is an example of the Sabbatian policy after Israel was established of migrating insider families to America so their children would be born American citizens. 'Obama' chose this financial team throughout his administration to respond to the Sabbatian-instigated crisis:

Timothy Geithner (ultra-Zionist) Treasury Secretary; Jacob J. Lew, Treasury Secretary; Larry Summers (ultra-Zionist), director of the White House National Economic Council; Paul Adolph Volcker (Rothschild business partner), chairman of the Economic Recovery Advisory Board; Peter Orszag (ultra-Zionist), director of the Office of Management and Budget overseeing all government spending; Penny Pritzker (ultra-Zionist), Commerce Secretary; Jared Bernstein (ultra-Zionist), chief economist and economic policy adviser to Vice President Joe Biden; Mary Schapiro (ultra-Zionist), chair of the Securities and Exchange Commission (SEC); Gary Gensler (ultra-Zionist), chairman of the Commodity Futures Trading Commission (CFTC); Sheila Bair (ultra-Zionist), chair of the Federal Deposit Insurance Corporation (FDIC); Karen Mills (ultra-Zionist), head of the Small Business Administration (SBA); Kenneth Feinberg (ultra-Zionist), Special Master for Executive [bail-out] Compensation. Feinberg would be appointed to oversee compensation (with strings) to 9/11 victims and families in a campaign to stop them having their day in court to question the official story. At the same time ultra-Zionist Bernard Bernanke was chairman of the Federal Reserve and these are only some of the ultra-Zionists with allegiance to Sabbatian-controlled Israel in the Obama government. Obama's biggest corporate donor was ultra-Zionist Goldman Sachs which had employed many in his administration.

Donald Trump ('Republican')

Trump claimed to be an outsider (he wasn't) who had come to 'drain the swamp'. He embarked on this goal by immediately appointing ultra-Zionist Steve Mnuchin, a Goldman Sachs employee for 17

years, as his Treasury Secretary. Others included Gary Cohn (ultra-Zionist), chief operating officer of Goldman Sachs, his first Director of the National Economic Council and chief economic adviser, who was later replaced by Larry Kudlow (ultra-Zionist). Trump's senior adviser throughout his four years in the White House was his sinister son-in-law Jared Kushner, a life-long friend of Israel Prime Minister Benjamin Netanyahu. Kushner is the son of a convicted crook who was pardoned by Trump in his last days in office. Other ultra-Zionists in the Trump administration included: Stephen Miller, Senior Policy Adviser; Avrahm Berkowitz, Deputy Adviser to Trump and his Senior Adviser Jared Kushner; Ivanka Trump, Adviser to the President, who converted to Judaism when she married Jared Kushner; David Friedman, Trump lawyer and Ambassador to Israel; Jason Greenblatt, Trump Organization executive vice president and chief legal officer, who was made Special Representative for International Negotiations and the Israeli-Palestinian Conflict; Rod Rosenstein, Deputy Attorney General; Elliot Abrams, Special Representative for Venezuela, then Iran; John Eisenberg, National Security Council Legal Adviser and Deputy Council to the President for National Security Affairs; Anne Neuberger, Deputy National Manager, National Security Agency; Ezra Cohen-Watnick, Acting Under Secretary of Defense for Intelligence; Elan Carr, Special Envoy to monitor and combat anti-Semitism; Len Khodorkovsky, Deputy Special Envoy to monitor and combat anti-Semitism; Reed Cordish, Assistant to the President, Intragovernmental and Technology Initiatives. Trump Vice President Mike Pence and Secretary of State Mike Pompeo, both Christian Zionists, were also vehement supporters of Israel and its goals and ambitions.

Donald 'free-speech believer' Trump pardoned a number of financial and violent criminals while ignoring calls to pardon Julian Assange and Edward Snowden whose crimes are revealing highly relevant information about government manipulation and corruption and the widespread illegal surveillance of the American people by US 'security' agencies. It's so good to know that Trump is on the side of freedom and justice and not mega-criminals with

allegiance to Sabbatian-controlled Israel. These included a pardon for Israeli spy Jonathan Pollard who was jailed for life in 1987 under the Espionage Act. Aviem Sella, the Mossad agent who recruited Pollard, was also pardoned by Trump while Assange sat in jail and Snowden remained in exile in Russia. Sella had 'fled' (was helped to escape) to Israel in 1987 and was never extradited despite being charged under the Espionage Act. A Trump White House statement said that Sella's clemency had been 'supported by Benjamin Netanyahu, Ron Dermer, Israel's US Ambassador, David Friedman, US Ambassador to Israel and Miriam Adelson, wife of leading Trump donor Sheldon Adelson who died shortly before. Other friends of Jared Kushner were pardoned along with Sholom Weiss who was believed to be serving the longest-ever white-collar prison sentence of more than 800 years in 2000. The sentence was commuted of Ponzi-schemer Eliyahu Weinstein who defrauded Jews and others out of $200 million. I did mention that Assange and Snowden were ignored, right? Trump gave Sabbatians almost everything they asked for in military and political support, moving the US Embassy from Tel Aviv to Jerusalem with its critical symbolic and literal implications for Palestinian statehood, and the 'deal of the Century' designed by Jared Kushner and David Friedman which gave the Sabbatian Israeli government the green light to substantially expand its already widespread program of building illegal Jewish-only settlements in the occupied land of the West Bank. This made a two-state 'solution' impossible by seizing all the land of a potential Palestinian homeland and that had been the plan since 1948 and then 1967 when the Arab-controlled Gaza Strip, West Bank, Sinai Peninsula and Syrian Golan Heights were occupied by Israel. All the talks about talks and road maps and delays have been buying time until the West Bank was physically occupied by Israeli real estate. Trump would have to be a monumentally ill-informed idiot not to see that this was the plan he was helping to complete. The Trump administration was in so many ways the Kushner administration which means the Netanyahu administration which means the Sabbatian administration. I understand why many opposing Cult fascism in all its forms gravitated to Trump, but he

was a crucial part of the Sabbatian plan and I will deal with this in the next chapter.

Joe Biden ('Democrat')

A barely cognitive Joe Biden took over the presidency in January, 2021, along with his fellow empty shell, Vice-President Kamala Harris, as the latest Sabbatian gofers to enter the White House. Names on the door may have changed and the 'party' – the force behind them remained the same as Zionists were appointed to a stream of pivotal areas relating to Sabbatian plans and policy. They included: Janet Yellen, Treasury Secretary, former head of the Federal Reserve, and still another ultra-Zionist running the US Treasury after Mnuchin (Trump), Lew and Geithner (Obama), and Summers and Rubin (Clinton); Anthony Blinken, Secretary of State; Wendy Sherman, Deputy Secretary of State (so that's 'Biden's' Sabbatian foreign policy sorted); Jeff Zients, White House coronavirus coordinator; Rochelle Walensky, head of the Centers for Disease Control; Rachel Levine, transgender deputy health secretary (that's 'Covid' hoax policy under control); Merrick Garland, Attorney General; Alejandro Mayorkas, Secretary of Homeland Security; Cass Sunstein, Homeland Security with responsibility for new immigration laws; Avril Haines, Director of National Intelligence; Anne Neuberger, National Security Agency cybersecurity director (note, cybersecurity); David Cohen, CIA Deputy Director; Ronald Klain, Biden's Chief of Staff (see Rahm Emanuel); Eric Lander, a 'leading geneticist', Office of Science and Technology Policy director (see Smart Grid, synthetic biology agenda); Jessica Rosenworcel, acting head of the Federal Communications Commission (FCC) which controls Smart Grid technology policy and electromagnetic communication systems including 5G. How can it be that so many pivotal positions are held by two-percent of the American population and 0.2 percent of the world population administration after administration no matter who is the president and what is the party? It's a coincidence? Of course it's not and this is why Sabbatians have built their colossal global web of interlocking 'anti-

hate' hate groups to condemn anyone who asks these glaring questions as an 'anti-Semite'. The way that Jewish people horrifically abused in Sabbatian-backed Nazi Germany are exploited to this end is stomach-turning and disgusting beyond words.

Political fusion

Sabbatian manipulation has reversed the roles of Republicans and Democrats and the same has happened in Britain with the Conservative and Labour Parties. Republicans and Conservatives were always labelled the 'right' and Democrats and Labour the 'left', but look at the policy positions now and the Democrat-Labour 'left' has moved further to the 'right' than Republicans and Conservatives under the banner of 'Woke', the Cult-created far-right tyranny. Where once the Democrat-Labour 'left' defended free speech and human rights they now seek to delete them and as I said earlier despite the 'Covid' fascism of the Jackboot Johnson Conservative government in the UK the Labour Party of leader Keir Starmer demanded even more extreme measures. The Labour Party has been very publicly absorbed by Sabbatians after a political and media onslaught against the previous leader, the weak and inept Jeremy Corbyn, over made-up allegations of 'anti-Semitism' both by him and his party. The plan was clear with this 'anti-Semite' propaganda and what was required in response was a swift and decisive 'fuck off' from Corbyn and a statement to expose the Anti-Semitism Industry (Sabbatian) attempt to silence Labour criticism of the Israeli government (Sabbatians) and purge the party of all dissent against the extremes of ultra-Zionism (Sabbatians). Instead Corbyn and his party fell to their knees and appeased the abusers which, by definition, is impossible. Appeasing one demand leads only to a new demand to be appeased until takeover is complete. Like I say – 'fuck off' would have been a much more effective policy and I have used it myself with great effect over the years when Sabbatians are on my case which is most of the time. I consider that fact a great compliment, by the way. The outcome of the Labour Party capitulation is that we now have a Sabbatian-controlled

Conservative Party 'opposed' by a Sabbatian-controlled Labour Party in a one-party Sabbatian state that hurtles towards the extremes of tyranny (the Sabbatian cult agenda). In America the situation is the same. Labour's Keir Starmer spends his days on his knees with his tongue out pointing to Tel Aviv, or I guess now Jerusalem, while Boris Johnson has an 'anti-Semitism czar' in the form of former Labour MP John Mann who keeps Starmer company on his prayer mat.

Sabbatian influence can be seen in Jewish members of the Labour Party who have been ejected for criticism of Israel including those from families that suffered in Nazi Germany. Sabbatians despise real Jewish people and target them even more harshly because it is so much more difficult to dub them 'anti-Semitic' although in their desperation they do try.

CHAPTER THREE

The Pushbacker sting

Until you realize how easy it is for your mind to be manipulated, you remain the puppet of someone else's game
Evita Ochel

I will use the presidencies of Trump and Biden to show how the manipulation of the one-party state plays out behind the illusion of political choice across the world. No two presidencies could – on the face of it – be more different and apparently at odds in terms of direction and policy.

A Renegade Mind sees beyond the obvious and focuses on outcomes and consequences and not image, words and waffle. The Cult embarked on a campaign to divide America between those who blindly support its agenda (the mentality known as 'Woke') and those who are pushing back on where the Cult and its Sabbatians want to go. This presents infinite possibilities for dividing and ruling the population by setting them at war with each other and allows a perceptual ring fence of demonisation to encircle the Pushbackers in a modern version of the Little Big Horn in 1876 when American cavalry led by Lieutenant Colonel George Custer were drawn into a trap, surrounded and killed by Native American tribes defending their land of thousands of years from being seized by the government. In this modern version the roles are reversed and it's those defending themselves from the Sabbatian government who are surrounded and the government that's seeking to destroy them. This trap was set years ago and to explain how we must return to 2016

and the emergence of Donald Trump as a candidate to be President of the United States. He set out to overcome the best part of 20 other candidates in the Republican Party before and during the primaries and was not considered by many in those early stages to have a prayer of living in the White House. The Republican Party was said to have great reservations about Trump and yet somehow he won the nomination. When you know how American politics works – politics in general – there is no way that Trump could have become the party's candidate unless the Sabbatian-controlled 'Neocons' that run the Republican Party wanted that to happen. We saw the proof in emails and documents made public by WikiLeaks that the Democratic Party hierarchy, or Democons, systematically undermined the campaign of Bernie Sanders to make sure that Sabbatian gofer Hillary Clinton won the nomination to be their presidential candidate. If the Democons could do that then the Neocons in the Republican Party could have derailed Trump in the same way. But they didn't and at that stage I began to conclude that Trump could well be the one chosen to be president. If that was the case the 'why' was pretty clear to see – the goal of dividing America between Cult agenda-supporting Wokers and Pushbackers who gravitated to Trump because he was telling them what they wanted to hear. His constituency of support had been increasingly ignored and voiceless for decades and profoundly through the eight years of Sabbatian puppet Barack Obama. Now here was someone speaking their language of pulling back from the incessant globalisation of political and economic power, the exporting of American jobs to China and elsewhere by 'American' (Sabbatian) corporations, the deletion of free speech, and the mass immigration policies that had further devastated job opportunities for the urban working class of all races and the once American heartlands of the Midwest.

Beware the forked tongue

Those people collectively sighed with relief that at last a political leader was apparently on their side, but another trait of the Renegade Mind is that you look even harder at people telling you

what you want to hear than those who are telling you otherwise. Obviously as I said earlier people wish what they want to hear to be true and genuine and they are much more likely to believe that than someone saying what they don't want to here and don't want to be true. Sales people are taught to be skilled in eliciting by calculated questioning what their customers want to hear and repeating that back to them as their own opinion to get their targets to like and trust them. Assets of the Cult are also sales people in the sense of selling perception. To read Cult manipulation you have to play the long and expanded game and not fall for the Vaudeville show of party politics. Both American parties are vehicles for the Cult and they exploit them in different ways depending on what the agenda requires at that moment. Trump and the Republicans were used to be the focus of dividing America and isolating Pushbackers to open the way for a Biden presidency to become the most extreme in American history by advancing the full-blown Woke (Cult) agenda with the aim of destroying and silencing Pushbackers now labelled Nazi Trump supporters and white supremacists.

Sabbatians wanted Trump in office for the reasons described by ultra-Zionist Saul Alinsky (1909-1972) who was promoting the Woke philosophy through 'community organising' long before anyone had heard of it. In those days it still went by its traditional name of Marxism. The reason for the manipulated Trump phenomenon was laid out in Alinsky's 1971 book, *Rules for Radicals*, which was his blueprint for overthrowing democratic and other regimes and replacing them with Sabbatian Marxism. Not surprisingly his to-do list was evident in the Sabbatian French and Russian 'Revolutions' and that in China which will become very relevant in the next chapter about the 'Covid' hoax. Among Alinsky's followers have been the deeply corrupt Barack Obama, House Speaker Nancy Pelosi and Hillary Clinton who described him as a 'hero'. All three are Sabbatian stooges with Pelosi personifying the arrogant corrupt idiocy that so widely fronts up for the Cult inner core. Predictably as a Sabbatian advocate of the 'light-bringer' Alinsky features Lucifer on the dedication page of his book as the original radical who gained

his own kingdom ('Earth' as we shall see). One of Alinsky's golden radical rules was to pick an individual and focus all attention, hatred and blame on them and not to target faceless bureaucracies and corporations. *Rules for Radicals* is really a Sabbatian handbook with its contents repeatedly employed all over the world for centuries and why wouldn't Sabbatians bring to power their designer-villain to be used as the individual on which all attention, hatred and blame was bestowed? This is what they did and the only question for me is how much Trump knew that and how much he was manipulated. A bit of both, I suspect. This was Alinsky's Trump technique from a man who died in 1972. The technique has spanned history:

Pick the target, freeze it, personalize it, polarize it. Don't try to attack abstract corporations or bureaucracies. Identify a responsible individual. Ignore attempts to shift or spread the blame.

From the moment Trump came to illusory power everything was about him. It wasn't about Republican policy or opinion, but all about Trump. Everything he did was presented in negative, derogatory and abusive terms by the Sabbatian-dominated media led by Cult operations such as CNN, MSNBC, *The New York Times* and the Jeff Bezos-owned *Washington Post* – 'Pick the target, freeze it, personalize it, polarize it.' Trump was turned into a demon to be vilified by those who hated him and a demi-god loved by those who worshipped him. This, in turn, had his supporters, too, presented as equally demonic in preparation for the punchline later down the line when Biden was about to take office. It was here's a Trump, there's a Trump, everywhere a Trump, Trump. Virtually every news story or happening was filtered through the lens of 'The Donald'. You loved him or hated him and which one you chose was said to define you as Satan's spawn or a paragon of virtue. Even supporting some Trump policies or statements and not others was enough for an assault on your character. No shades of grey were or are allowed. Everything is black and white (literally and figuratively). A Californian I knew had her head utterly scrambled by her hatred for Trump while telling people they should love each other. She was so totally consumed by

Trump Derangement Syndrome as it became to be known that this glaring contradiction would never have occurred to her. By definition anyone who criticised Trump or praised his opponents was a hero and this lady described Joe Biden as 'a kind, honest gentleman' when he's a provable liar, mega-crook and vicious piece of work to boot. Sabbatians had indeed divided America using Trump as the fall-guy and all along the clock was ticking on the consequences for his supporters.

In hock to his masters

Trump gave Sabbatians via Israel almost everything they wanted in his four years. Ask and you shall receive was the dynamic between himself and Benjamin Netanyahu orchestrated by Trump's ultra-Zionist son-in-law Jared Kushner, his ultra-Zionist Ambassador to Israel, David Friedman, and ultra-Zionist 'Israel adviser', Jason Greenblatt. The last two were central to the running and protecting from collapse of his business empire, the Trump Organisation, and colossal business failures made him forever beholding to Sabbatian networks that bailed him out. By the start of the 1990s Trump owed $4 billion to banks that he couldn't pay and almost $1billion of that was down to him personally and not his companies. This mega-disaster was the result of building two new casinos in Atlantic City and buying the enormous Taj Mahal operation which led to crippling debt payments. He had borrowed fantastic sums from 72 banks with major Sabbatian connections and although the scale of debt should have had him living in a tent alongside the highway they never foreclosed. A plan was devised to lift Trump from the mire by BT Securities Corporation and Rothschild Inc. and the case was handled by Wilber Ross who had worked for the Rothschilds for 27 years. Ross would be named US Commerce Secretary after Trump's election. Another crucial figure in saving Trump was ultra-Zionist 'investor' Carl Icahn who bought the Taj Mahal casino. Icahn was made special economic adviser on financial regulation in the Trump administration. He didn't stay long but still managed to find time to make a tidy sum of a reported $31.3 million when he sold his

holdings affected by the price of steel three days before Trump imposed a 235 percent tariff on steel imports. What amazing bits of luck these people have. Trump and Sabbatian operatives have long had a close association and his mentor and legal adviser from the early 1970s until 1986 was the dark and genetically corrupt ultra-Zionist Roy Cohn who was chief counsel to Senator Joseph McCarthy's 'communist' witch-hunt in the 1950s. *Esquire* magazine published an article about Cohn with the headline 'Don't mess with Roy Cohn'. He was described as the most feared lawyer in New York and 'a ruthless master of dirty tricks ... [with] ... more than one Mafia Don on speed dial'. Cohn's influence, contacts, support and protection made Trump a front man for Sabbatians in New York with their connections to one of Cohn's many criminal employers, the 'Russian' Sabbatian Mafia. Israel-centric media mogul Rupert Murdoch was introduced to Trump by Cohn and they started a long friendship. Cohn died in 1986 weeks after being disbarred for unethical conduct by the Appellate Division of the New York State Supreme Court. The wheels of justice do indeed run slow given the length of Cohn's crooked career.

QAnon-sense

We are asked to believe that Donald Trump with his fundamental connections to Sabbatian networks and operatives has been leading the fight to stop the Sabbatian agenda for the fascistic control of America and the world. Sure he has. A man entrapped during his years in the White House by Sabbatian operatives and whose biggest financial donor was casino billionaire Sheldon Adelson who was Sabbatian to his DNA?? Oh, do come on. Trump has been used to divide America and isolate Pushbackers on the Cult agenda under the heading of 'Trump supporters', 'insurrectionists' and 'white supremacists'. The US Intelligence/Mossad Psyop or psychological operation known as QAnon emerged during the Trump years as a central pillar in the Sabbatian campaign to lead Pushbackers into the trap set by those that wished to destroy them. I knew from the start that QAnon was a scam because I had seen the same scenario many

times before over 30 years under different names and I had written about one in particular in the books. 'Not again' was my reaction when QAnon came to the fore. The same script is pulled out every few years and a new name added to the letterhead. The story always takes the same form: 'Insiders' or 'the good guys' in the government-intelligence-military 'Deep State' apparatus were going to instigate mass arrests of the 'bad guys' which would include the Rockefellers, Rothschilds, Barack Obama, Hillary Clinton, George Soros, etc., etc. Dates are given for when the 'good guys' are going to move in, but the dates pass without incident and new dates are given which pass without incident. The central message to Pushbackers in each case is that they don't have to do anything because there is 'a plan' and it is all going to be sorted by the 'good guys' on the inside. 'Trust the plan' was a QAnon mantra when the only plan was to misdirect Pushbackers into putting their trust in a Psyop they believed to be real. Beware, beware, those who tell you what you want to hear and always check it out. Right up to Biden's inauguration QAnon was still claiming that 'the Storm' was coming and Trump would stay on as president when Biden and his cronies were arrested and jailed. It was never going to happen and of course it didn't, but what did happen as a result provided that punchline to the Sabbatian Trump/QAnon Psyop.

On January 6th, 2021, a very big crowd of Trump supporters gathered in the National Mall in Washington DC down from the Capitol Building to protest at what they believed to be widespread corruption and vote fraud that stopped Trump being re-elected for a second term as president in November, 2020. I say as someone that does not support Trump or Biden that the evidence is clear that major vote-fixing went on to favour Biden, a man with cognitive problems so advanced he can often hardly string a sentence together without reading the words written for him on the Teleprompter. Glaring ballot discrepancies included serious questions about electronic voting machines that make vote rigging a comparative cinch and hundreds of thousands of paper votes that suddenly appeared during already advanced vote counts and virtually all of

them for Biden. Early Trump leads in crucial swing states suddenly began to close and disappear. The pandemic hoax was used as the excuse to issue almost limitless numbers of mail-in ballots with no checks to establish that the recipients were still alive or lived at that address. They were sent to streams of people who had not even asked for them. Private organisations were employed to gather these ballots and who knows what they did with them before they turned up at the counts. The American election system has been manipulated over decades to become a sick joke with more holes than a Swiss cheese for the express purpose of dictating the results. Then there was the criminal manipulation of information by Sabbatian tech giants like Facebook, Twitter and Google-owned YouTube which deleted pro-Trump, anti-Biden accounts and posts while everything in support of Biden was left alone. Sabbatians wanted Biden to win because after the dividing of America it was time for full-on Woke and every aspect of the Cult agenda to be unleashed.

Hunter gatherer

Extreme Silicon Valley bias included blocking information by the *New York Post* exposing a Biden scandal that should have ended his bid for president in the final weeks of the campaign. Hunter Biden, his monumentally corrupt son, is reported to have sent a laptop to be repaired at a local store and failed to return for it. Time passed until the laptop became the property of the store for non-payment of the bill. When the owner saw what was on the hard drive he gave a copy to the FBI who did nothing even though it confirmed widespread corruption in which the Joe Biden family were using his political position, especially when he was vice president to Obama, to make multiple millions in countries around the world and most notably Ukraine and China. Hunter Biden's one-time business partner Tony Bobulinski went public when the story broke in the *New York Post* to confirm the corruption he saw and that Joe Biden not only knew what was going on he also profited from the spoils. Millions were handed over by a Chinese company with close

connections – like all major businesses in China – to the Chinese communist party of President Xi Jinping. Joe Biden even boasted at a meeting of the Cult's World Economic Forum that as vice president he had ordered the government of Ukraine to fire a prosecutor. What he didn't mention was that the same man just happened to be investigating an energy company which was part of Hunter Biden's corrupt portfolio. The company was paying him big bucks for no other reason than the influence his father had. Overnight Biden's presidential campaign should have been over given that he had lied publicly about not knowing what his son was doing. Instead almost the entire Sabbatian-owned mainstream media and Sabbatian-owned Silicon Valley suppressed circulation of the story. This alone went a mighty way to rigging the election of 2020. Cult assets like Mark Zuckerberg at Facebook also spent hundreds of millions to be used in support of Biden and vote 'administration'.

The Cult had used Trump as the focus to divide America and was now desperate to bring in moronic, pliable, corrupt Biden to complete the double-whammy. No way were they going to let little things like the will of the people thwart their plan. Silicon Valley widely censored claims that the election was rigged because it *was* rigged. For the same reason anyone claiming it was rigged was denounced as a 'white supremacist' including the pathetically few Republican politicians willing to say so. Right across the media where the claim was mentioned it was described as a 'false claim' even though these excuses for 'journalists' would have done no research into the subject whatsoever. Trump won seven million more votes than any sitting president had ever achieved while somehow a cognitively-challenged soon to be 78-year-old who was hidden away from the public for most of the campaign managed to win more votes than any presidential candidate in history. It makes no sense. You only had to see election rallies for both candidates to witness the enthusiasm for Trump and the apathy for Biden. Tens of thousands would attend Trump events while Biden was speaking in empty car parks with often only television crews attending and framing their shots to hide the fact that no one was there. It was pathetic to see

footage come to light of Biden standing at a podium making speeches only to TV crews and party fixers while reading the words written for him on massive Teleprompter screens. So, yes, those protestors on January 6th had a point about election rigging, but some were about to walk into a trap laid for them in Washington by the Cult Deep State and its QAnon Psyop. This was the Capitol Hill riot ludicrously dubbed an 'insurrection'.

The spider and the fly

Renegade Minds know there are not two 'sides' in politics, only one side, the Cult, working through all 'sides'. It's a stage show, a puppet show, to direct the perceptions of the population into focusing on diversions like parties and candidates while missing the puppeteers with their hands holding all the strings. The Capitol Hill 'insurrection' brings us back to the Little Big Horn. Having created two distinct opposing groupings – Woke and Pushbackers – the trap was about to be sprung. Pushbackers were to be encircled and isolated by associating them all in the public mind with Trump and then labelling Trump as some sort of Confederate leader. I knew immediately that the Capitol riot was a set-up because of two things. One was how easy the rioters got into the building with virtually no credible resistance and secondly I could see – as with the 'Covid' hoax in the West at the start of 2020 – how the Cult could exploit the situation to move its agenda forward with great speed. My experience of Cult techniques and activities over more than 30 years has showed me that while they do exploit situations they haven't themselves created this never happens with events of fundamental agenda significance. Every time major events giving cultists the excuse to rapidly advance their plan you find they are manipulated into being for the specific reason of providing that excuse – Problem-Reaction-Solution. Only a tiny minority of the huge crowd of Washington protestors sought to gain entry to the Capitol by smashing windows and breaching doors. That didn't matter. The whole crowd and all Pushbackers, even if they did not support Trump, were going to be lumped together as dangerous

insurrectionists and conspiracy theorists. The latter term came into widespread use through a CIA memo in the 1960s aimed at discrediting those questioning the nonsensical official story of the Kennedy assassination and it subsequently became widely employed by the media. It's still being used by inept 'journalists' with no idea of its origin to discredit anyone questioning anything that authority claims to be true. When you are perpetrating a conspiracy you need to discredit the very word itself even though the dictionary definition of conspiracy is merely 'the activity of secretly planning with other people to do something bad or illegal' and 'a general agreement to keep silent about a subject for the purpose of keeping it secret'. On that basis there are conspiracies almost wherever you look. For obvious reasons the Cult and its lapdog media have to claim there are no conspiracies even though the word appears in state laws as with conspiracy to defraud, to murder, and to corrupt public morals.

Agent provocateurs are widely used by the Cult Deep State to manipulate genuine people into acting in ways that suit the desired outcome. By genuine in this case I mean protestors genuinely supporting Trump and claims that the election was stolen. In among them, however, were agents of the state wearing the garb of Trump supporters and QAnon to pump-prime the Capital riot which some genuine Trump supporters naively fell for. I described the situation as 'Come into my parlour said the spider to the fly'. Leaflets appeared through the Woke paramilitary arm Antifa, the anti-fascist fascists, calling on supporters to turn up in Washington looking like Trump supporters even though they hated him. Some of those arrested for breaching the Capitol Building were sourced to Antifa and its stable mate Black Lives Matter. Both organisations are funded by Cult billionaires and corporations. One man charged for the riot was according to his lawyer a former FBI agent who had held top secret security clearance for 40 years. Attorney Thomas Plofchan said of his client, 66-year-old Thomas Edward Caldwell:

He has held a Top Secret Security Clearance since 1979 and has undergone multiple Special Background Investigations in support of his clearances. After retiring from the Navy, he

worked as a section chief for the Federal Bureau of Investigation from 2009-2010 as a GS-12 [mid-level employee].

He also formed and operated a consulting firm performing work, often classified, for U.S government customers including the US. Drug Enforcement Agency, Department of Housing and Urban Development, the US Coast Guard, and the US Army Personnel Command.

A judge later released Caldwell pending trial in the absence of evidence about a conspiracy or that he tried to force his way into the building. *The New York Post* reported a 'law enforcement source' as saying that 'at least two known Antifa members were spotted' on camera among Trump supporters during the riot while one of the rioters arrested was John Earle Sullivan, a seriously extreme Black Lives Matter Trump-hater from Utah who was previously arrested and charged in July, 2020, over a BLM-Antifa riot in which drivers were threatened and one was shot. Sullivan is the founder of Utah-based Insurgence USA which is an affiliate of the Cult-created-and-funded Black Lives Matter movement. Footage appeared and was then deleted by Twitter of Trump supporters calling out Antifa infiltrators and a group was filmed changing into pro-Trump clothing before the riot. Security at the building was *pathetic* – as planned. Colonel Leroy Fletcher Prouty, a man with long experience in covert operations working with the US security apparatus, once described the tell-tale sign to identify who is involved in an assassination. He said:

No one has to direct an assassination – it happens. The active role is played secretly by permitting it to happen. This is the greatest single clue. Who has the power to call off or reduce the usual security precautions?

This principle applies to many other situations and certainly to the Capitol riot of January 6th, 2021.

The sting

With such a big and potentially angry crowd known to be gathering near the Capitol the security apparatus would have had a major police detail to defend the building with National Guard troops on

standby given the strength of feeling among people arriving from all over America encouraged by the QAnon Psyop and statements by Donald Trump. Instead Capitol Police 'security' was flimsy, weak, and easily breached. The same number of officers was deployed as on a regular day and that is a blatant red flag. They were not staffed or equipped for a possible riot that had been an obvious possibility in the circumstances. No protective and effective fencing worth the name was put in place and there were no contingency plans. The whole thing was basically a case of standing aside and waving people in. Once inside police mostly backed off apart from one Capitol police officer who ridiculously shot dead unarmed Air Force veteran protestor Ashli Babbitt without a warning as she climbed through a broken window. The 'investigation' refused to name or charge the officer after what must surely be considered a murder in the circumstances. They just lifted a carpet and swept. The story was endlessly repeated about five people dying in the 'armed insurrection' when there was no report of rioters using weapons. Apart from Babbitt the other four died from a heart attack, strokes and apparently a drug overdose. Capitol police officer Brian Sicknick was reported to have died after being bludgeoned with a fire extinguisher when he was alive after the riot was over and died later of what the Washington Medical Examiner's Office said was a stroke. Sicknick had no external injuries. The lies were delivered like rapid fire. There was a narrative to build with incessant repetition of the lie until the lie became the accepted 'everybody knows that' truth. The 'Big Lie' technique of Nazi Propaganda Minister Joseph Goebbels is constantly used by the Cult which was behind the Nazis and is today behind the 'Covid' and 'climate change' hoaxes. Goebbels said:

If you tell a lie big enough and keep repeating it, people will eventually come to believe it. The lie can be maintained only for such time as the State can shield the people from the political, economic and/or military consequences of the lie. It thus becomes vitally important for the State to use all of its powers to repress dissent, for the truth is the mortal enemy of the lie, and thus by extension, the truth is the greatest enemy of the State.

Most protestors had a free run of the Capitol Building. This allowed pictures to be taken of rioters in iconic parts of the building including the Senate chamber which could be used as propaganda images against all Pushbackers. One Congresswoman described the scene as 'the worst kind of non-security anybody could ever imagine'. Well, the first part was true, but someone obviously did imagine it and made sure it happened. Some photographs most widely circulated featured people wearing QAnon symbols and now the Psyop would be used to dub all QAnon followers with the ubiquitous fit-all label of 'white supremacist' and 'insurrectionists'. When a Muslim extremist called Noah Green drove his car at two police officers at the Capitol Building killing one in April, 2021, there was no such political and media hysteria. They were just disappointed he wasn't white.

The witch-hunt

Government prosecutor Michael Sherwin, an aggressive, dark-eyed, professional Rottweiler led the 'investigation' and to call it over the top would be to understate reality a thousand fold. Hundreds were tracked down and arrested for the crime of having the wrong political views and people were jailed who had done nothing more than walk in the building, committed no violence or damage to property, took a few pictures and left. They were labelled a 'threat to the Republic' while Biden sat in the White House signing executive orders written for him that were dismantling 'the Republic'. Even when judges ruled that a mother and son should not be in jail the government kept them there. Some of those arrested have been badly beaten by prison guards in Washington and lawyers for one man said he suffered a fractured skull and was made blind in one eye. Meanwhile a woman is shot dead for no reason by a Capitol Police officer and we are not allowed to know who he is never mind what has happened to him although that will be *nothing*. The Cult's QAnon/Trump sting to identify and isolate Pushbackers and then target them on the road to crushing and deleting them was a resounding success. You would have thought the Russians had

invaded the building at gunpoint and lined up senators for a firing squad to see the political and media reaction. Congresswoman Alexandria Ocasio-Cortez is a child in a woman's body, a terrible-twos, me, me, me, Woker narcissist of such proportions that words have no meaning. She said she thought she was going to die when 'insurrectionists' banged on her office door. It turned out she wasn't even in the Capitol Building when the riot was happening and the 'banging' was a Capitol Police officer. She referred to herself as a 'survivor' which is an insult to all those true survivors of violent and sexual abuse while she lives her pampered and privileged life talking drivel for a living. Her Woke colleague and fellow mega-narcissist Rashida Tlaib broke down describing the devastating effect on her, too, of *not being* in the building when the rioters were there. Ocasio-Cortez and Tlaib are members of a fully-Woke group of Congresswomen known as 'The Squad' along with Ilhan Omar and Ayanna Pressley. The Squad from what I can see can be identified by its vehement anti-white racism, anti-white men agenda, and, as always in these cases, the absence of brain cells on active duty.

The usual suspects were on the riot case immediately in the form of Democrat ultra-Zionist senators and operatives Chuck Schumer and Adam Schiff demanding that Trump be impeached for 'his part in the insurrection'. The same pair of prats had led the failed impeachment of Trump over the invented 'Russia collusion' nonsense which claimed Russia had helped Trump win the 2016 election. I didn't realise that Tel Aviv had been relocated just outside Moscow. I must find an up-to-date map. The Russia hoax was a Sabbatian operation to keep Trump occupied and impotent and to stop any rapport with Russia which the Cult wants to retain as a perceptual enemy to be pulled out at will. Puppet Biden began attacking Russia when he came to office as the Cult seeks more upheaval, division and war across the world. A two-year stage show 'Russia collusion inquiry' headed by the not-very-bright former 9/11 FBI chief Robert Mueller, with support from 19 lawyers, 40 FBI agents plus intelligence analysts, forensic accountants and other

staff, devoured tens of millions of dollars and found no evidence of Russia collusion which a ten-year-old could have told them on day one. Now the same moronic Schumer and Schiff wanted a second impeachment of Trump over the Capitol 'insurrection' (riot) which the arrested development of Schumer called another 'Pearl Harbor' while others compared it with 9/11 in which 3,000 died and, in the case of CNN, with the Rwandan genocide in the 1990s in which an estimated 500,000 to 600,000 were murdered, between 250, 000 and 500,000 women were raped, and populations of whole towns were hacked to death with machetes. To make those comparisons purely for Cult political reasons is beyond insulting to those that suffered and lost their lives and confirms yet again the callous inhumanity that we are dealing with. Schumer is a monumental idiot and so is Schiff, but they serve the Cult agenda and do whatever they're told so they get looked after. Talking of idiots – another inane man who spanned the Russia and Capitol impeachment attempts was Senator Eric Swalwell who had the nerve to accuse Trump of collusion with the Russians while sleeping with a Chinese spy called Christine Fang or 'Fang Fang' which is straight out of a Bond film no doubt starring Klaus Schwab as the bloke living on a secret island and controlling laser weapons positioned in space and pointing at world capitals. Fang Fang plays the part of Bond's infiltrator girlfriend which I'm sure she would enjoy rather more than sharing a bed with the brainless Swalwell, lying back and thinking of China. The FBI eventually warned Swalwell about Fang Fang which gave her time to escape back to the Chinese dictatorship. How very thoughtful of them. The second Trump impeachment also failed and hardly surprising when an impeachment is supposed to remove a sitting president and by the time it happened Trump was no longer president. These people are running your country America, well, officially anyway. Terrifying isn't it?

Outcomes tell the story - always

The outcome of all this – and it's the *outcome* on which Renegade Minds focus, not the words – was that a vicious, hysterical and

obviously pre-planned assault was launched on Pushbackers to censor, silence and discredit them and even targeted their right to earn a living. They have since been condemned as 'domestic terrorists' that need to be treated like Al-Qaeda and Islamic State. 'Domestic terrorists' is a label the Cult has been trying to make stick since the period of the Oklahoma bombing in 1995 which was blamed on 'far-right domestic terrorists'. If you read *The Trigger* you will see that the bombing was clearly a Problem-Reaction-Solution carried out by the Deep State during a Bill Clinton administration so corrupt that no dictionary definition of the term would even nearly suffice. Nearly 30, 000 troops were deployed from all over America to the empty streets of Washington for Biden's inauguration. Ten thousand of them stayed on with the pretext of protecting the capital from insurrectionists when it was more psychological programming to normalise the use of the military in domestic law enforcement in support of the Cult plan for a police-military state. Biden's fascist administration began a purge of 'wrong-thinkers' in the military which means anyone that is not on board with Woke. The Capitol Building was surrounded by a fence with razor wire and the Land of the Free was further symbolically and literally dismantled. The circle was completed with the installation of Biden and the exploitation of the QAnon Psyop.

America had never been so divided since the civil war of the 19th century, Pushbackers were isolated and dubbed terrorists and now, as was always going to happen, the Cult immediately set about deleting what little was left of freedom and transforming American society through a swish of the hand of the most controlled 'president' in American history leading (officially at least) the most extreme regime since the country was declared an independent state on July 4th, 1776. Biden issued undebated, dictatorial executive orders almost by the hour in his opening days in office across the whole spectrum of the Cult wish-list including diluting controls on the border with Mexico allowing thousands of migrants to illegally enter the United States to transform the demographics of America and import an election-changing number of perceived Democrat

voters. Then there were Biden deportation amnesties for the already illegally resident (estimated to be as high as 20 or even 30 million). A bill before Congress awarded American citizenship to anyone who could prove they had worked in agriculture for just 180 days in the previous two years as 'Big Ag' secured its slave labour long-term. There were the plans to add new states to the union such as Puerto Rico and making Washington DC a state. They are all parts of a plan to ensure that the Cult-owned Woke Democrats would be permanently in power.

Border – what border?

I have exposed in detail in other books how mass immigration into the United States and Europe is the work of Cult networks fuelled by the tens of billions spent to this and other ends by George Soros and his global Open Society (open borders) Foundations. The impact can be seen in America alone where the population has increased by *100 million* in little more than 30 years mostly through immigration. I wrote in *The Answer* that the plan was to have so many people crossing the southern border that the numbers become unstoppable and we are now there under Cult-owned Biden. El Salvador in Central America puts the scale of what is happening into context. A third of the population now lives in the United States, much of it illegally, and many more are on the way. The methodology is to crush Central and South American countries economically and spread violence through machete-wielding psychopathic gangs like MS-13 based in El Salvador and now operating in many American cities. Biden-imposed lax security at the southern border means that it is all but open. He said before his 'election' that he wanted to see a surge towards the border if he became president and that was the green light for people to do just that after election day to create the human disaster that followed for both America and the migrants. When that surge came the imbecilic Alexandria Ocasio-Cortez said it wasn't a 'surge' because they are 'children, not insurgents' and the term 'surge' (used by Biden) was a claim of 'white supremacists'.

This disingenuous lady may one day enter the realm of the most basic intelligence, but it won't be any time soon.

Sabbatians and the Cult are in the process of destroying America by importing violent people and gangs in among the genuine to terrorise American cities and by overwhelming services that cannot cope with the sheer volume of new arrivals. Something similar is happening in Europe as Western society in general is targeted for demographic and cultural transformation and upheaval. The plan demands violence and crime to create an environment of intimidation, fear and division and Soros has been funding the election of district attorneys across America who then stop prosecuting many crimes, reduce sentences for violent crimes and free as many violent criminals as they can. Sabbatians are creating the chaos from which order – their order – can respond in a classic Problem-Reaction-Solution. A Freemasonic moto says 'Ordo Ab Chao' (Order out of Chaos) and this is why the Cult is constantly creating chaos to impose a new 'order'. Here you have the reason the Cult is constantly creating chaos. The 'Covid' hoax can be seen with those entering the United States by plane being forced to take a 'Covid' test while migrants flooding through southern border processing facilities do not. Nothing is put in the way of mass migration and if that means ignoring the government's own 'Covid' rules then so be it. They know it's all bullshit anyway. Any pushback on this is denounced as 'racist' by Wokers and Sabbatian fronts like the ultra-Zionist Anti-Defamation League headed by the appalling Jonathan Greenblatt which at the same time argues that Israel should not give citizenship and voting rights to more Palestinian Arabs or the 'Jewish population' (in truth the Sabbatian network) will lose control of the country.

Society-changing numbers

Biden's masters have declared that countries like El Salvador are so dangerous that their people must be allowed into the United States for humanitarian reasons when there are fewer murders in large parts of many Central American countries than in US cities like

Baltimore. That is not to say Central America cannot be a dangerous place and Cult-controlled American governments have been making it so since way back, along with the dismantling of economies, in a long-term plan to drive people north into the United States. Parts of Central America are very dangerous, but in other areas the story is being greatly exaggerated to justify relaxing immigration criteria. Migrants are being offered free healthcare and education in the United States as another incentive to head for the border and there is no requirement to be financially independent before you can enter to prevent the resources of America being drained. You can't blame migrants for seeking what they believe will be a better life, but they are being played by the Cult for dark and nefarious ends. The numbers since Biden took office are huge. In February, 2021, more than 100,000 people were known to have tried to enter the US illegally through the southern border (it was 34,000 in the same month in 2020) and in March it was 170,000 – a 418 percent increase on March, 2020. These numbers are only known people, not the ones who get in unseen. The true figure for migrants illegally crossing the border in a single month was estimated by one congressman at 250,000 and that number will only rise under Biden's current policy. Gangs of murdering drug-running thugs that control the Mexican side of the border demand money – thousands of dollars – to let migrants cross the Rio Grande into America. At the same time gun battles are breaking out on the border several times a week between rival Mexican drug gangs (which now operate globally) who are equipped with sophisticated military-grade weapons, grenades and armoured vehicles. While the Capitol Building was being 'protected' from a non-existent 'threat' by thousands of troops, and others were still deployed at the time in the Cult Neocon war in Afghanistan, the southern border of America was left to its fate. This is not incompetence, it is cold calculation.

By March, 2021, there were 17,000 unaccompanied children held at border facilities and many of them are ensnared by people traffickers for paedophile rings and raped on their journey north to America. This is not conjecture – this is fact. Many of those designated

children are in reality teenage boys or older. Meanwhile Wokers posture their self-purity for encouraging poor and tragic people to come to America and face this nightmare both on the journey and at the border with the disgusting figure of House Speaker Nancy Pelosi giving disingenuous speeches about caring for migrants. The woman's evil. Wokers condemned Trump for having children in cages at the border (so did Obama, *Shhhh*), but now they are sleeping on the floor without access to a shower with one border facility 729 percent over capacity. The Biden insanity even proposed flying migrants from the southern border to the northern border with Canada for 'processing'. The whole shambles is being overseen by ultra-Zionist Secretary of Homeland Security, the moronic liar Alejandro Mayorkas, who banned news cameras at border facilities to stop Americans seeing what was happening. Mayorkas said there was not a ban on news crews; it was just that they were not allowed to film. Alongside him at Homeland Security is another ultra-Zionist Cass Sunstein appointed by Biden to oversee new immigration laws. Sunstein despises conspiracy researchers to the point where he suggests they should be banned or *taxed* for having such views. The man is not bonkers or anything. He's perfectly well-adjusted, but adjusted to what is the question. Criticise what is happening and you are a 'white supremacist' when earlier non-white immigrants also oppose the numbers which effect their lives and opportunities. Black people in poor areas are particularly damaged by uncontrolled immigration and the increased competition for work opportunities with those who will work for less. They are also losing voting power as Hispanics become more dominant in former black areas. It's a downward spiral for them while the billionaires behind the policy drone on about how much they care about black people and 'racism'. None of this is about compassion for migrants or black people – that's just wind and air. Migrants are instead being mercilessly exploited to transform America while the countries they leave are losing their future and the same is true in Europe. Mass immigration may now be the work of Woke Democrats, but it can be traced back to the 1986 Immigration Reform and Control Act (it

wasn't) signed into law by Republican hero President Ronald Reagan which gave amnesty to millions living in the United States illegally and other incentives for people to head for the southern border. Here we have the one-party state at work again.

Save me syndrome

Almost every aspect of what I have been exposing as the Cult agenda was on display in even the first days of 'Biden' with silencing of Pushbackers at the forefront of everything. A Renegade Mind will view the Trump years and QAnon in a very different light to their supporters and advocates as the dots are connected. The QAnon/Trump Psyop has given the Cult all it was looking for. We may not know how much, or little, that Trump realised he was being used, but that's a side issue. This pincer movement produced the desired outcome of dividing America and having Pushbackers isolated. To turn this around we have to look at new routes to empowerment which do not include handing our power to other people and groups through what I will call the 'Save Me Syndrome' – 'I want someone else to do it so that I don't have to'. We have seen this at work throughout human history and the QAnon/Trump Psyop is only the latest incarnation alongside all the others. Religion is an obvious expression of this when people look to a 'god' or priest to save them or tell them how to be saved and then there are 'save me' politicians like Trump. Politics is a diversion and not a 'saviour'. It is a means to block positive change, not make it possible.

Save Me Syndrome always comes with the same repeating theme of handing your power to whom or what you believe will save you while your real 'saviour' stares back from the mirror every morning. Renegade Minds are constantly vigilant in this regard and always asking the question 'What can I do?' rather than 'What can someone else do for me?' Gandhi was right when he said: 'You must be the change you want to see in the world.' We are indeed the people we have been waiting for. We are presented with a constant raft of reasons to concede that power to others and forget where the real power is. Humanity has the numbers and the Cult does not. It has to

use diversion and division to target the unstoppable power that comes from unity. Religions, governments, politicians, corporations, media, QAnon, are all different manifestations of this power-diversion and dilution. Refusing to give your power to governments and instead handing it to Trump and QAnon is not to take a new direction, but merely to recycle the old one with new names on the posters. I will explore this phenomenon as we proceed and how to break the cycles and recycles that got us here through the mists of repeating perception and so repeating history.

For now we shall turn to the most potent example in the entire human story of the consequences that follow when you give your power away. I am talking, of course, of the 'Covid' hoax.

CHAPTER FOUR

'Covid': Calculated catastrophe

Facts are threatening to those invested in fraud
DaShanne Stokes

We can easily unravel the real reason for the 'Covid pandemic' hoax by employing the Renegade Mind methodology that I have outlined this far. We'll start by comparing the long-planned Cult outcome with the 'Covid pandemic' outcome. Know the outcome and you'll see the journey.

I have highlighted the plan for the Hunger Games Society which has been in my books for so many years with the very few controlling the very many through ongoing dependency. To create this dependency it is essential to destroy independent livelihoods, businesses and employment to make the population reliant on the state (the Cult) for even the basics of life through a guaranteed pittance income. While independence of income remained these Cult ambitions would be thwarted. With this knowledge it was easy to see where the 'pandemic' hoax was going once talk of 'lockdowns' began and the closing of all but perceived 'essential' businesses to 'save' us from an alleged 'deadly virus'. Cult corporations like Amazon and Walmart were naturally considered 'essential' while mom and pop shops and stores had their doors closed by fascist decree. As a result with every new lockdown and new regulation more small and medium, even large businesses not owned by the Cult, went to the wall while Cult giants and their frontmen and women grew financially fatter by the second. Mom and pop were

denied an income and the right to earn a living and the wealth of people like Jeff Bezos (Amazon), Mark Zuckerberg (Facebook) and Sergei Brin and Larry Page (Google/Alphabet) have reached record levels. The Cult was increasing its own power through further dramatic concentrations of wealth while the competition was being destroyed and brought into a state of dependency. Lockdowns have been instigated to secure that very end and were never anything to do with health. My brother Paul spent 45 years building up a bus repair business, but lockdowns meant buses were running at a fraction of normal levels for months on end. Similar stories can told in their hundreds of millions worldwide. Efforts of a lifetime coldly destroyed by Cult multi-billionaires and their lackeys in government and law enforcement who continued to earn their living from the taxation of the people while denying the right of the same people to earn theirs. How different it would have been if those making and enforcing these decisions had to face the same financial hardships of those they affected, but they never do.

Gates of Hell

Behind it all in the full knowledge of what he is doing and why is the psychopathic figure of Cult operative Bill Gates. His puppet Tedros at the World Health Organization declared 'Covid' a pandemic in March, 2020. The WHO had changed the definition of a 'pandemic' in 2009 just a month before declaring the 'swine flu pandemic' which would not have been so under the previous definition. The same applies to 'Covid'. The definition had included… 'an infection by an infectious agent, occurring simultaneously in different countries, with a significant mortality rate relative to the proportion of the population infected'. The new definition removed the need for 'significant mortality'. The 'pandemic' has been fraudulent even down to the definition, but Gates demanded economy-destroying lockdowns, school closures, social distancing, mandatory masks, a 'vaccination' for every man, woman and child on the planet and severe consequences and restrictions for those that refused. Who gave him this power? The

Cult did which he serves like a little boy in short trousers doing what his daddy tells him. He and his psychopathic missus even smiled when they said that much worse was to come (what they knew was planned to come). Gates responded in the matter-of-fact way of all psychopaths to a question about the effect on the world economy of what he was doing:

> Well, it won't go to zero but it will shrink. Global GDP is probably going to take the biggest hit ever [Gates was smiling as he said this] … in my lifetime this will be the greatest economic hit. But you don't have a choice. People act as if you have a choice. People don't feel like going to the stadium when they might get infected … People are deeply affected by seeing these stats, by knowing they could be part of the transmission chain, old people, their parents and grandparents, could be affected by this, and so you don't get to say ignore what is going on here.
>
> There will be the ability to open up, particularly in rich countries, if things are done well over the next few months, but for the world at large normalcy only returns when we have largely vaccinated the entire population.

The man has no compassion or empathy. How could he when he's a psychopath like all Cult players? My own view is that even beyond that he is very seriously mentally ill. Look in his eyes and you can see this along with his crazy flailing arms. You don't do what he has done to the world population since the start of 2020 unless you are mentally ill and at the most extreme end of psychopathic. You especially don't do it when to you know, as we shall see, that cases and deaths from 'Covid' are fakery and a product of monumental figure massaging. 'These stats' that Gates referred to are based on a 'test' that's not testing for the 'virus' as he has known all along. He made his fortune with big Cult support as an infamously ruthless software salesman and now buys global control of 'health' (death) policy without the population he affects having any say. It's a breathtaking outrage. Gates talked about people being deeply affected by fear of 'Covid' when that was because of *him* and his global network lying to them minute-by-minute supported by a lying media that he seriously influences and funds to the tune of hundreds of millions. He's handed big sums to media operations including the BBC, NBC, Al Jazeera, Univision, *PBS NewsHour*,

ProPublica, National Journal, The Guardian, The Financial Times, The Atlantic, Texas Tribune, USA Today publisher Gannett, *Washington Monthly, Le Monde,* Center for Investigative Reporting, Pulitzer Center on Crisis Reporting, National Press Foundation, International Center for Journalists, Solutions Journalism Network, the Poynter Institute for Media Studies, and many more. Gates is everywhere in the 'Covid' hoax and the man must go to prison – or a mental facility – for the rest of his life and his money distributed to those he has taken such enormous psychopathic pleasure in crushing.

The Muscle

The Hunger Games global structure demands a police-military state – a fusion of the two into one force – which viciously imposes the will of the Cult on the population and protects the Cult from public rebellion. In that regard, too, the 'Covid' hoax just keeps on giving. Often unlawful, ridiculous and contradictory 'Covid' rules and regulations have been policed across the world by moronic automatons and psychopaths made faceless by face-nappy masks and acting like the Nazi SS and fascist blackshirts and brownshirts of Hitler and Mussolini. The smallest departure from the rules decreed by the psychos in government and their clueless gofers were jumped upon by the face-nappy fascists. Brutality against public protestors soon became commonplace even on girls, women and old people as the brave men with the batons – the Face-Nappies as I call them – broke up peaceful protests and handed out fines like confetti to people who couldn't earn a living let alone pay hundreds of pounds for what was once an accepted human right. Robot Face-Nappies of Nottingham police in the English East Midlands fined one group £11,000 for attending a child's birthday party. For decades I charted the transformation of law enforcement as genuine, decent officers were replaced with psychopaths and the brain dead who would happily and brutally do whatever their masters told them. Now they were let loose on the public and I would emphasise the point that none of this just happened. The step-by-step change in the dynamic between police and public was orchestrated from the shadows by

those who knew where this was all going and the same with the perceptual reframing of those in all levels of authority and official administration through 'training courses' by organisations such as Common Purpose which was created in the late 1980s and given a massive boost in Blair era Britain until it became a global phenomenon. Supposed public 'servants' began to view the population as the enemy and the same was true of the police. This was the start of the explosion of behaviour manipulation organisations and networks preparing for the all-war on the human psyche unleashed with the dawn of 2020. I will go into more detail about this later in the book because it is a core part of what is happening.

Police desecrated beauty spots to deter people gathering and arrested women for walking in the countryside alone 'too far' from their homes. We had arrogant, clueless sergeants in the Isle of Wight police where I live posting on Facebook what they insisted the population must do or else. A schoolmaster sergeant called Radford looked young enough for me to ask if his mother knew he was out, but he was posting what he *expected* people to do while a Sergeant Wilkinson boasted about fining lads for meeting in a McDonald's car park where they went to get a lockdown takeaway. Wilkinson added that he had even cancelled their order. What a pair of prats these people are and yet they have increasingly become the norm among Jackboot Johnson's Yellowshirts once known as the British police. This was the theme all over the world with police savagery common during lockdown protests in the United States, the Netherlands, and the fascist state of Victoria in Australia under its tyrannical and again moronic premier Daniel Andrews. Amazing how tyrannical and moronic tend to work as a team and the same combination could be seen across America as arrogant, narcissistic Woke governors and mayors such as Gavin Newsom (California), Andrew Cuomo (New York), Gretchen Whitmer (Michigan), Lori Lightfoot (Chicago) and Eric Garcetti (Los Angeles) did their Nazi and Stalin impressions with the full support of the compliant brutality of their enforcers in uniform as they arrested small business owners defying

fascist shutdown orders and took them to jail in ankle shackles and handcuffs. This happened to bistro owner Marlena Pavlos-Hackney in Gretchen Whitmer's fascist state of Michigan when police arrived to enforce an order by a state-owned judge for 'putting the community at risk' at a time when other states like Texas were dropping restrictions and migrants were pouring across the southern border without any 'Covid' questions at all. I'm sure there are many officers appalled by what they are ordered to do, but not nearly enough of them. If they were truly appalled they would not do it. As the months passed every opportunity was taken to have the military involved to make their presence on the streets ever more familiar and 'normal' for the longer-term goal of police-military fusion.

Another crucial element to the Hunger Games enforcement network has been encouraging the public to report neighbours and others for 'breaking the lockdown rules'. The group faced with £11,000 in fines at the child's birthday party would have been dobbed-in by a neighbour with a brain the size of a pea. The technique was most famously employed by the Stasi secret police in communist East Germany who had public informants placed throughout the population. A police chief in the UK says his force doesn't need to carry out 'Covid' patrols when they are flooded with so many calls from the public reporting other people for visiting the beach. Dorset police chief James Vaughan said people were so enthusiastic about snitching on their fellow humans they were now operating as an auxiliary arm of the police: 'We are still getting around 400 reports a week from the public, so we will respond to reports …We won't need to be doing hotspot patrols because people are very quick to pick the phone up and tell us.' Vaughan didn't say that this is a pillar of all tyrannies of whatever complexion and the means to hugely extend the reach of enforcement while spreading distrust among the people and making them wary of doing anything that might get them reported. Those narcissistic Isle of Wight sergeants Radford and Wilkinson never fail to add a link to their Facebook posts where the public can inform on their fellow slaves.

Neither would be self-aware enough to realise they were imitating the Stasi which they might well never have heard of. Government psychologists that I will expose later laid out a policy to turn communities against each other in the same way.

A coincidence? Yep, and I can knit fog

I knew from the start of the alleged pandemic that this was a Cult operation. It presented limitless potential to rapidly advance the Cult agenda and exploit manipulated fear to demand that every man, woman and child on the planet was 'vaccinated' in a process never used on humans before which infuses self-replicating *synthetic* material into human cells. Remember the plan to transform the human body from a biological to a synthetic biological state. I'll deal with the 'vaccine' (that's not actually a vaccine) when I focus on the genetic agenda. Enough to say here that mass global 'vaccination' justified by this 'new virus' set alarms ringing after 30 years of tracking these people and their methods. The 'Covid' hoax officially beginning in China was also a big red flag for reasons I will be explaining. The agenda potential was so enormous that I could dismiss any idea that the 'virus' appeared naturally. Major happenings with major agenda implications never occur without Cult involvement in making them happen. My questions were twofold in early 2020 as the media began its campaign to induce global fear and hysteria: Was this alleged infectious agent released on purpose by the Cult or did it even exist at all? I then did what I always do in these situations. I sat, observed and waited to see where the evidence and information would take me. By March and early April synchronicity was strongly – and ever more so since then – pointing me in the direction of *there is no 'virus'*. I went public on that with derision even from swathes of the alternative media that voiced a scenario that the Chinese government released the 'virus' in league with Deep State elements in the United States from a top-level bio-lab in Wuhan where the 'virus' is said to have first appeared. I looked at that possibility, but I didn't buy it for several reasons. Deaths from the 'virus' did not in any way match what they

would have been with a 'deadly bioweapon' and it is much more effective if you sell the *illusion* of an infectious agent rather than having a real one unless you can control through injection who has it and who doesn't. Otherwise you lose control of events. A made-up 'virus' gives you a blank sheet of paper on which you can make it do whatever you like and have any symptoms or mutant 'variants' you choose to add while a real infectious agent would limit you to what it actually does. A phantom disease allows you to have endless ludicrous 'studies' on the 'Covid' dollar to widen the perceived impact by inventing ever more 'at risk' groups including one study which said those who walk slowly may be almost four times more likely to die from the 'virus'. People are in psychiatric wards for less.

A real 'deadly bioweapon' can take out people in the hierarchy that are not part of the Cult, but essential to its operation. Obviously they don't want that. Releasing a real disease means you immediately lose control of it. Releasing an illusory one means you don't. Again it's vital that people are extra careful when dealing with what they want to hear. A bioweapon unleashed from a Chinese laboratory in collusion with the American Deep State may fit a conspiracy narrative, but is it true? Would it not be far more effective to use the excuse of a 'virus' to justify the real bioweapon – the 'vaccine'? That way your disease agent does not have to be transmitted and arrives directly through a syringe. I saw a French virologist Luc Montagnier quoted in the alternative media as saying he had discovered that the alleged 'new' severe acute respiratory syndrome coronavirus , or SARS-CoV-2, was made artificially and included elements of the human immunodeficiency 'virus' (HIV) and a parasite that causes malaria. SARS-CoV-2 is alleged to trigger an alleged illness called Covid-19. I remembered Montagnier's name from my research years before into claims that an HIV 'retrovirus' causes AIDs – claims that were demolished by Berkeley virologist Peter Duesberg who showed that no one had ever proved that HIV causes acquired immunodeficiency syndrome or AIDS. Claims that become accepted as fact, publicly and medically, with no proof whatsoever are an ever-recurring story that profoundly applies to

'Covid'. Nevertheless, despite the lack of proof, Montagnier's team at the Pasteur Institute in Paris had a long dispute with American researcher Robert Gallo over which of them discovered and isolated the HIV 'virus' and with *no evidence* found it to cause AIDS. You will see later that there is also no evidence that any 'virus' causes any disease or that there is even such a thing as a 'virus' in the way it is said to exist. The claim to have 'isolated' the HIV 'virus' will be presented in its real context as we come to the shocking story – and it is a story – of SARS-CoV-2 and so will Montagnier's assertion that he identified the full SARS-CoV-2 genome.

Hoax in the making

We can pick up the 'Covid' story in 2010 and the publication by the Rockefeller Foundation of a document called 'Scenarios for the Future of Technology and International Development'. The inner circle of the Rockefeller family has been serving the Cult since John D. Rockefeller (1839-1937) made his fortune with Standard Oil. It is less well known that the same Rockefeller – the Bill Gates of his day – was responsible for establishing what is now referred to as 'Big Pharma', the global network of pharmaceutical companies that make outrageous profits dispensing scalpel and drug 'medicine' and are obsessed with pumping vaccines in ever-increasing number into as many human arms and backsides as possible. John D. Rockefeller was the driving force behind the creation of the 'education' system in the United States and elsewhere specifically designed to program the perceptions of generations thereafter. The Rockefeller family donated exceptionally valuable land in New York for the United Nations building and were central in establishing the World Health Organization in 1948 as an agency of the UN which was created from the start as a Trojan horse and stalking horse for world government. Now enter Bill Gates. His family and the Rockefellers have long been extremely close and I have seen genealogy which claims that if you go back far enough the two families fuse into the same bloodline. Gates has said that the Bill and Melinda Gates Foundation was inspired by the Rockefeller Foundation and why not

when both are serving the same Cult? Major tax-exempt foundations are overwhelmingly criminal enterprises in which Cult assets fund the Cult agenda in the guise of 'philanthropy' while avoiding tax in the process. Cult operatives can become mega-rich in their role of front men and women for the psychopaths at the inner core and they, too, have to be psychopaths to knowingly serve such evil. Part of the deal is that a big percentage of the wealth gleaned from representing the Cult has to be spent advancing the ambitions of the Cult and hence you have the Rockefeller Foundation, Bill and Melinda Gates Foundation (and *so* many more) and people like George Soros with his global Open Society Foundations spending their billions in pursuit of global Cult control. Gates is a global public face of the Cult with his interventions in world affairs including Big Tech influence; a central role in the 'Covid' and 'vaccine' scam; promotion of the climate change shakedown; manipulation of education; geoengineering of the skies; and his food-control agenda as the biggest owner of farmland in America, his GMO promotion and through other means. As one writer said: 'Gates monopolizes or wields disproportionate influence over the tech industry, global health and vaccines, agriculture and food policy (including biopiracy and fake food), weather modification and other climate technologies, surveillance, education and media.' The almost limitless wealth secured through Microsoft and other not-allowed-to-fail ventures (including vaccines) has been ploughed into a long, long list of Cult projects designed to enslave the entire human race. Gates and the Rockefellers have been working as one unit with the Rockefeller-established World Health Organization leading global 'Covid' policy controlled by Gates through his mouth-piece Tedros. Gates became the WHO's biggest funder when Trump announced that the American government would cease its donations, but Biden immediately said he would restore the money when he took office in January, 2021. The Gates Foundation (the Cult) owns through limitless funding the world health system and the major players across the globe in the 'Covid' hoax.

Okay, with that background we return to that Rockefeller Foundation document of 2010 headed 'Scenarios for the Future of Technology and International Development' and its 'imaginary' epidemic of a virulent and deadly influenza strain which infected 20 percent of the global population and killed eight million in seven months. The Rockefeller scenario was that the epidemic destroyed economies, closed shops, offices and other businesses and led to governments imposing fierce rules and restrictions that included mandatory wearing of face masks and body-temperature checks to enter communal spaces like railway stations and supermarkets. The document predicted that even after the height of the Rockefeller-envisaged epidemic the authoritarian rule would continue to deal with further pandemics, transnational terrorism, environmental crises and rising poverty. Now you may think that the Rockefellers are our modern-day seers or alternatively, and rather more likely, that they well knew what was planned a few years further on. Fascism had to be imposed, you see, to 'protect citizens from risk and exposure'. The Rockefeller scenario document said:

During the pandemic, national leaders around the world flexed their authority and imposed airtight rules and restrictions, from the mandatory wearing of face masks to body-temperature checks at the entries to communal spaces like train stations and supermarkets. Even after the pandemic faded, this more authoritarian control and oversight of citizens and their activities stuck and even intensified. In order to protect themselves from the spread of increasingly global problems – from pandemics and transnational terrorism to environmental crises and rising poverty – leaders around the world took a firmer grip on power.

At first, the notion of a more controlled world gained wide acceptance and approval. Citizens willingly gave up some of their sovereignty – and their privacy – to more paternalistic states in exchange for greater safety and stability. Citizens were more tolerant, and even eager, for top-down direction and oversight, and national leaders had more latitude to impose order in the ways they saw fit.

In developed countries, this heightened oversight took many forms: biometric IDs for all citizens, for example, and tighter regulation of key industries whose stability was deemed vital to national interests. In many developed countries, enforced cooperation with a suite of new regulations and agreements slowly but steadily restored both order and, importantly, economic growth.

There we have the prophetic Rockefellers in 2010 and three years later came their paper for the Global Health Summit in Beijing, China, when government representatives, the private sector, international organisations and groups met to discuss the next 100 years of 'global health'. The Rockefeller Foundation-funded paper was called 'Dreaming the Future of Health for the Next 100 Years and more prophecy ensued as it described a dystopian future: 'The abundance of data, digitally tracking and linking people may mean the 'death of privacy' and may replace physical interaction with transient, virtual connection, generating isolation and raising questions of how values are shaped in virtual networks.' Next in the 'Covid' hoax preparation sequence came a 'table top' simulation in 2018 for another 'imaginary' pandemic of a disease called Clade X which was said to kill 900 million people. The exercise was organised by the Gates-funded Johns Hopkins University's Center for Health Security in the United States and this is the very same university that has been compiling the disgustingly and systematically erroneous global figures for 'Covid' cases and deaths. Similar Johns Hopkins health crisis scenarios have included the Dark Winter exercise in 2001 and Atlantic Storm in 2005.

Nostradamus 201

For sheer predictive genius look no further prophecy-watchers than the Bill Gates-funded Event 201 held only six weeks before the 'coronavirus pandemic' is supposed to have broken out in China and Event 201 was based on a scenario of a global 'coronavirus pandemic'. Melinda Gates, the great man's missus, told the BBC that he had 'prepared for years' for a coronavirus pandemic which told us what we already knew. Nostradamugates had predicted in a TED talk in 2015 that a pandemic was coming that would kill a lot of people and demolish the world economy. My god, the man is a machine – possibly even literally. Now here he was only weeks before the real thing funding just such a simulated scenario and involving his friends and associates at Johns Hopkins, the World Economic Forum Cult-front of Klaus Schwab, the United Nations,

Johnson & Johnson, major banks, and officials from China and the Centers for Disease Control in the United States. What synchronicity – Johns Hopkins would go on to compile the fraudulent 'Covid' figures, the World Economic Forum and Schwab would push the 'Great Reset' in response to 'Covid', the Centers for Disease Control would be at the forefront of 'Covid' policy in the United States, Johnson & Johnson would produce a 'Covid vaccine', and everything would officially start just weeks later in China. Spooky, eh? They were even accurate in creating a simulation of a 'virus' pandemic because the 'real thing' would also be a simulation. Event 201 was not an exercise preparing for something that might happen; it was a rehearsal for what those in control knew was *going* to happen and very shortly. Hours of this simulation were posted on the Internet and the various themes and responses mirrored what would soon be imposed to transform human society. News stories were inserted and what they said would be commonplace a few weeks later with still more prophecy perfection. Much discussion focused on the need to deal with misinformation and the 'anti-vax movement' which is exactly what happened when the 'virus' arrived – was said to have arrived – in the West.

Cult-owned social media banned criticism and exposure of the official 'virus' narrative and when I said there *was* no 'virus' in early April, 2020, I was banned by one platform after another including YouTube, Facebook and later Twitter. The mainstream broadcast media in Britain was in effect banned from interviewing me by the Tony-Blair-created government broadcasting censor Ofcom headed by career government bureaucrat Melanie Dawes who was appointed just as the 'virus' hoax was about to play out in January, 2020. At the same time the Ickonic media platform was using Vimeo, another ultra-Zionist-owned operation, while our own player was being created and they deleted in an instant hundreds of videos, documentaries, series and shows to confirm their unbelievable vindictiveness. We had copies, of course, and they had to be restored one by one when our player was ready. These people have no class. Sabbatian Facebook promised free advertisements for the Gates-

controlled World Health Organization narrative while deleting 'false claims and conspiracy theories' to stop 'misinformation' about the alleged coronavirus. All these responses could be seen just a short while earlier in the scenarios of Event 201. Extreme censorship was absolutely crucial for the Cult because the official story was so ridiculous and unsupportable by the evidence that it could never survive open debate and the free-flow of information and opinion. If you can't win a debate then don't have one is the Cult's approach throughout history. Facebook's little boy front man – front boy – Mark Zuckerberg equated 'credible and accurate information' with official sources and exposing their lies with 'misinformation'.

Silencing those that can see

The censorship dynamic of Event 201 is now the norm with an army of narrative-supporting 'fact-checker' organisations whose entire reason for being is to tell the public that official narratives are true and those exposing them are lying. One of the most appalling of these 'fact-checkers' is called NewsGuard founded by ultra-Zionist Americans Gordon Crovitz and Steven Brill. Crovitz is a former publisher of *The Wall Street Journal*, former Executive Vice President of Dow Jones, a member of the Council on Foreign Relations (CFR), and on the board of the American Association of Rhodes Scholars. The CFR and Rhodes Scholarships, named after Rothschild agent Cecil Rhodes who plundered the gold and diamonds of South Africa for his masters and the Cult, have featured widely in my books. NewsGuard don't seem to like me for some reason – I really can't think why – and they have done all they can to have me censored and discredited which is, to quote an old British politician, like being savaged by a dead sheep. They are, however, like all in the censorship network, very well connected and funded by organisations themselves funded by, or connected to, Bill Gates. As you would expect with anything associated with Gates NewsGuard has an offshoot called HealthGuard which 'fights online health care hoaxes'. How very kind. Somehow the NewsGuard European Managing Director Anna-Sophie Harling, a remarkably young-

looking woman with no broadcasting experience and little hands-on work in journalism, has somehow secured a position on the 'Content Board' of UK government broadcast censor Ofcom. An executive of an organisation seeking to discredit dissidents of the official narratives is making decisions for the government broadcast 'regulator' about content?? Another appalling 'fact-checker' is Full Fact funded by George Soros and global censors Google and Facebook.

It's amazing how many activists in the 'fact-checking', 'anti-hate', arena turn up in government-related positions – people like UK Labour Party activist Imran Ahmed who heads the Center for Countering Digital Hate founded by people like Morgan McSweeney, now chief of staff to the Labour Party's hapless and useless 'leader' Keir Starmer. Digital Hate – which is what it really is – uses the American spelling of Center to betray its connection to a transatlantic network of similar organisations which in 2020 shapeshifted from attacking people for 'hate' to attacking them for questioning the 'Covid' hoax and the dangers of the 'Covid vaccine'. It's just a coincidence, you understand. This is one of Imran Ahmed's hysterical statements: 'I would go beyond calling anti-vaxxers conspiracy theorists to say they are an extremist group that pose a national security risk.' No one could ever accuse this prat of understatement and he's including in that those parents who are now against vaccines after their children were damaged for life or killed by them. He's such a nice man. Ahmed does the rounds of the Woke media getting soft-ball questions from spineless 'journalists' who never ask what right he has to campaign to destroy the freedom of speech of others while he demands it for himself. There also seems to be an overrepresentation in Ofcom of people connected to the narrative-worshipping BBC. This incredible global network of narrative-support was super-vital when the 'Covid' hoax was played in the light of the mega-whopper lies that have to be defended from the spotlight cast by the most basic intelligence.

Setting the scene

The Cult plays the long game and proceeds step-by-step ensuring that everything is in place before major cards are played and they don't come any bigger than the 'Covid' hoax. The psychopaths can't handle events where the outcome isn't certain and as little as possible – preferably nothing – is left to chance. Politicians, government and medical officials who would follow direction were brought to illusory power in advance by the Cult web whether on the national stage or others like state governors and mayors of America. For decades the dynamic between officialdom, law enforcement and the public was changed from one of service to one of control and dictatorship. Behaviour manipulation networks established within government were waiting to impose the coming 'Covid' rules and regulations specifically designed to subdue and rewire the psyche of the people in the guise of protecting health. These included in the UK the Behavioural Insights Team part-owned by the British government Cabinet Office; the Scientific Pandemic Insights Group on Behaviours (SPI-B); and a whole web of intelligence and military groups seeking to direct the conversation on social media and control the narrative. Among them are the cyberwarfare (on the people) 77th Brigade of the British military which is also coordinated through the Cabinet Office as civilian and military leadership continues to combine in what they call the Fusion Doctrine. The 77th Brigade is a British equivalent of the infamous Israeli (Sabbatian) military cyberwarfare and Internet manipulation operation Unit 8200 which I expose at length in *The Trigger*. Also carefully in place were the medical and science advisers to government – many on the payroll past or present of Bill Gates – and a whole alternative structure of unelected government stood by to take control when elected parliaments were effectively closed down once the 'Covid' card was slammed on the table. The structure I have described here and so much more was installed in every major country through the Cult networks. The top-down control hierarchy looks like this: The Cult – Cult-owned Gates – the World Health Organization and Tedros – Gates-funded or controlled chief medical officers and science 'advisers' (dictators) in each country –

political 'leaders'– law enforcement – The People. Through this simple global communication and enforcement structure the policy of the Cult could be imposed on virtually the entire human population so long as they acquiesced to the fascism. With everything in place it was time for the button to be pressed in late 2019/early 2020.

These were the prime goals the Cult had to secure for its will to prevail:

1) Locking down economies, closing all but designated 'essential' businesses (Cult-owned corporations were 'essential'), and putting the population under house arrest was an imperative to destroy independent income and employment and ensure dependency on the Cult-controlled state in the Hunger Games Society. Lockdowns had to be established as the global blueprint from the start to respond to the 'virus' and followed by pretty much the entire world.

2) The global population had to be terrified into believing in a deadly 'virus' that didn't actually exist so they would unquestioningly obey authority in the belief that authority must know how best to protect them and their families. Software salesman Gates would suddenly morph into the world's health expert and be promoted as such by the Cult-owned media.

3) A method of testing that wasn't testing for the 'virus', but was only claimed to be, had to be in place to provide the illusion of 'cases' and subsequent 'deaths' that had a very different cause to the 'Covid-19' that would be scribbled on the death certificate.

4) Because there was no 'virus' and the great majority testing positive with a test not testing for the 'virus' would have no symptoms of anything the lie had to be sold that people without symptoms (without the 'virus') could still pass it on to others. This was crucial to justify for the first time quarantining – house arresting – healthy people. Without this the economy-destroying lockdown of *everybody* could not have been credibly sold.

5) The 'saviour' had to be seen as a vaccine which beyond evil drug companies were working like angels of mercy to develop as quickly as possible, with all corners cut, to save the day. The public must absolutely not know that the 'vaccine' had nothing to do with a 'virus' or that the contents were ready and waiting with a very different motive long before the 'Covid' card was even lifted from the pack.

I said in March, 2020, that the 'vaccine' would have been created way ahead of the 'Covid' hoax which justified its use and the following December an article in the New York *Intelligencer* magazine said the Moderna 'vaccine' had been 'designed' by

January, 2020. This was 'before China had even acknowledged that the disease could be transmitted from human to human, more than a week before the first confirmed coronavirus case in the United States'. The article said that by the time the first American death was announced a month later 'the vaccine had already been manufactured and shipped to the National Institutes of Health for the beginning of its Phase I clinical trial'. The 'vaccine' was actually 'designed' long before that although even with this timescale you would expect the article to ask how on earth it could have been done that quickly. Instead it asked why the 'vaccine' had not been rolled out then and not months later. Journalism in the mainstream is truly dead. I am going to detail in the next chapter why the 'virus' has never existed and how a hoax on that scale was possible, but first the foundation on which the Big Lie of 'Covid' was built.

The test that doesn't test

Fraudulent 'testing' is the bottom line of the whole 'Covid' hoax and was the means by which a 'virus' that did not exist *appeared* to exist. They could only achieve this magic trick by using a test not testing for the 'virus'. To use a test that *was* testing for the 'virus' would mean that every test would come back negative given there was no 'virus'. They chose to exploit something called the RT-PCR test invented by American biochemist Kary Mullis in the 1980s who said publicly that his PCR test … *cannot detect infectious disease*. Yes, the 'test' used worldwide to detect infectious 'Covid' to produce all the illusory 'cases' and 'deaths' compiled by Johns Hopkins and others *cannot detect infectious disease*. This fact came from the mouth of the man who invented PCR and was awarded the Nobel Prize in Chemistry in 1993 for doing so. Sadly, and incredibly conveniently for the Cult, Mullis died in August, 2019, at the age of 74 just before his test would be fraudulently used to unleash fascism on the world. He was said to have died from pneumonia which was an irony in itself. A few months later he would have had 'Covid-19' on his death certificate. I say the timing of his death was convenient because had he lived Mullis, a brilliant, honest and decent man, would have been

vociferously speaking out against the use of his test to detect 'Covid' when it was never designed, or able, to do that. I know that to be true given that Mullis made the same point when his test was used to 'detect' – not detect – HIV. He had been seriously critical of the Gallo/Montagnier claim to have isolated the HIV 'virus' and shown it to cause AIDS for which Mullis said there was no evidence. AIDS is actually not a disease but a series of diseases from which people die all the time. When they die from those *same diseases* after a positive 'test' for HIV then AIDS goes on their death certificate. I think I've heard that before somewhere. Countries instigated a policy with 'Covid' that anyone who tested positive with a test not testing for the 'virus' and died of any other cause within 28 days and even longer 'Covid-19' had to go on the death certificate. Cases have come from the test that can't test for infectious disease and the deaths are those who have died of *anything* after testing positive with a test not testing for the 'virus'. I'll have much more later about the death certificate scandal.

Mullis was deeply dismissive of the now US 'Covid' star Anthony Fauci who he said was a liar who didn't know anything about anything – 'and I would say that to his face – nothing.' He said of Fauci: 'The man thinks he can take a blood sample, put it in an electron microscope and if it's got a virus in there you'll know it – he doesn't understand electron microscopy and he doesn't understand medicine and shouldn't be in a position like he's in.' That position, terrifyingly, has made him the decider of 'Covid' fascism policy on behalf of the Cult in his role as director since 1984 of the National Institute of Allergy and Infectious Diseases (NIAID) while his record of being wrong is laughable; but being wrong, so long as it's the *right kind* of wrong, is why the Cult loves him. He'll say anything the Cult tells him to say. Fauci was made Chief Medical Adviser to the President immediately Biden took office. Biden was installed in the White House by Cult manipulation and one of his first decisions was to elevate Fauci to a position of even more control. This is a coincidence? Yes, and I identify as a flamenco dancer called Lola. How does such an incompetent criminal like Fauci remain in that

pivotal position in American health since *the 1980s*? When you serve the Cult it looks after you until you are surplus to requirements. Kary Mullis said prophetically of Fauci and his like: 'Those guys have an agenda and it's not an agenda we would like them to have … they make their own rules, they change them when they want to, and Tony Fauci does not mind going on television in front of the people who pay his salary and lie directly into the camera.' Fauci has done that almost daily since the 'Covid' hoax began. Lying is in Fauci's DNA. To make the situation crystal clear about the PCR test this is a direct quote from its inventor Kary Mullis:

> It [the PCR test] doesn't tell you that you're sick and doesn't tell you that the thing you ended up with was really going to hurt you ...'

Ask yourself why governments and medical systems the world over have been using this very test to decide who is 'infected' with the SARS-CoV-2 'virus' and the alleged disease it allegedly causes, 'Covid-19'. The answer to that question will tell you what has been going on. By the way, here's a little show-stopper – the 'new' SARS-CoV-2 'virus' was 'identified' as such right from the start using … *the PCR test not testing for the 'virus'*. If you are new to this and find that shocking then stick around. I have hardly started yet. Even worse, other 'tests', like the 'Lateral Flow Device' (LFD), are considered so useless that they have to be *confirmed* by the PCR test! Leaked emails written by Ben Dyson, adviser to UK 'Health' Secretary Matt Hancock, said they were 'dangerously unreliable'. Dyson, executive director of strategy at the Department of Health, wrote: 'As of today, someone who gets a positive LFD result in (say) London has at best a 25 per cent chance of it being a true positive, but if it is a self-reported test potentially as low as 10 per cent (on an optimistic assumption about specificity) or as low as 2 per cent (on a more pessimistic assumption).' These are the 'tests' that schoolchildren and the public are being urged to have twice a week or more and have to isolate if they get a positive. Each fake positive goes in the statistics as a 'case' no matter how ludicrously inaccurate and the

'cases' drive lockdown, masks and the pressure to 'vaccinate'. The government said in response to the email leak that the 'tests' were accurate which confirmed yet again what shocking bloody liars they are. The real false positive rate is *100 percent* as we'll see. In another 'you couldn't make it up' the UK government agreed to pay £2.8 billion to California's Innova Medical Group to supply the irrelevant lateral flow tests. The company's primary test-making centre is in China. Innova Medical Group, established in March, 2020, is owned by Pasaca Capital Inc, chaired by Chinese-American millionaire Charles Huang who was born in Wuhan.

How it works – and how it doesn't

The RT-PCR test, known by its full title of Polymerase chain reaction, is used across the world to make millions, even billions, of copies of a DNA/RNA genetic information sample. The process is called 'amplification' and means that a tiny sample of genetic material is amplified to bring out the detailed content. I stress that it is not testing for an infectious disease. It is simply amplifying a sample of genetic material. In the words of Kary Mullis: 'PCR is … just a process that's used to make a whole lot of something out of something.' To emphasise the point companies that make the PCR tests circulated around the world to 'test' for 'Covid' warn on the box that it can't be used to detect 'Covid' or infectious disease and is for research purposes only. It's okay, rest for a minute and you'll be fine. This is the test that produces the 'cases' and 'deaths' that have been used to destroy human society. All those global and national medical and scientific 'experts' demanding this destruction to 'save us' *KNOW* that the test is not testing for the 'virus' and the cases and deaths they claim to be real are an almost unimaginable fraud. Every one of them and so many others including politicians and psychopaths like Gates and Tedros must be brought before Nuremburg-type trials and jailed for the rest of their lives. The more the genetic sample is amplified by PCR the more elements of that material become sensitive to the test and by that I don't mean sensitive for a 'virus' but for elements of the genetic material which

is *naturally* in the body or relates to remnants of old conditions of various kinds lying dormant and causing no disease. Once the amplification of the PCR reaches a certain level *everyone* will test positive. So much of the material has been made sensitive to the test that everyone will have some part of it in their body. Even lying criminals like Fauci have said that once PCR amplifications pass 35 cycles everything will be a false positive that cannot be trusted for the reasons I have described. I say, like many proper doctors and scientists, that 100 percent of the 'positives' are false, but let's just go with Fauci for a moment.

He says that any amplification over 35 cycles will produce false positives and yet the US Centers for Disease Control (CDC) and Food and Drug Administration (FDA) have recommended up to *40 cycles* and the National Health Service (NHS) in Britain admitted in an internal document for staff that it was using *45 cycles* of amplification. A long list of other countries has been doing the same and at least one 'testing' laboratory has been using *50 cycles*. Have you ever heard a doctor, medical 'expert' or the media ask what level of amplification has been used to claim a 'positive'. The 'test' comes back 'positive' and so you have the 'virus', end of story. Now we can see how the government in Tanzania could send off samples from a goat and a pawpaw fruit under human names and both came back positive for 'Covid-19'. Tanzania president John Magufuli mocked the 'Covid' hysteria, the PCR test and masks and refused to import the DNA-manipulating 'vaccine'. The Cult hated him and an article sponsored by the Bill Gates Foundation appeared in the London *Guardian* in February, 2021, headed 'It's time for Africa to rein in Tanzania's anti-vaxxer president'. Well, 'reined in' he shortly was. Magufuli appeared in good health, but then, in March, 2021, he was dead at 61 from 'heart failure'. He was replaced by Samia Hassan Suhulu who is connected to Klaus Schwab's World Economic Forum and she immediately reversed Magufuli's 'Covid' policy. A sample of cola tested positive for 'Covid' with the PCR test in Germany while American actress and singer-songwriter Erykah Badu tested positive in one nostril and negative in the other. Footballer Ronaldo called

the PCR test 'bullshit' after testing positive three times and being forced to quarantine and miss matches when there was nothing wrong with him. The mantra from Tedros at the World Health Organization and national governments (same thing) has been test, test, test. They know that the more tests they can generate the more fake 'cases' they have which go on to become 'deaths' in ways I am coming to. The UK government has its Operation Moonshot planned to test multiple millions every day in workplaces and schools with free tests for everyone to use twice a week at home in line with the Cult plan from the start to make testing part of life. A government advertisement for an 'Interim Head of Asymptomatic Testing Communication' said the job included responsibility for delivering a 'communications strategy' (propaganda) 'to support the expansion of asymptomatic testing that *'normalises testing as part of everyday life*'. More tests means more fake 'cases', 'deaths' and fascism. I have heard of, and from, many people who booked a test, couldn't turn up, and yet got a positive result through the post for a test they'd never even had. The whole thing is crazy, but for the Cult there's method in the madness. Controlling and manipulating the level of amplification of the test means the authorities can control whenever they want the number of apparent 'cases' and 'deaths'. If they want to justify more fascist lockdown and destruction of livelihoods they keep the amplification high. If they want to give the illusion that lockdowns and the 'vaccine' are working then they lower the amplification and 'cases' and 'deaths' will appear to fall. In January, 2021, the Cult-owned World Health Organization suddenly warned laboratories about over-amplification of the test and to lower the threshold. Suddenly headlines began appearing such as: 'Why ARE "Covid" cases plummeting?' This was just when the vaccine rollout was underway and I had predicted months before they would make cases appear to fall through amplification tampering when the 'vaccine' came. These people are so predictable.

Cow vaccines?

The question must be asked of what is on the test swabs being poked far up the nose of the population to the base of the brain? A nasal swab punctured one woman's brain and caused it to leak fluid. Most of these procedures are being done by people with little training or medical knowledge. Dr Lorraine Day, former orthopaedic trauma surgeon and Chief of Orthopaedic Surgery at San Francisco General Hospital, says the tests are really a *'vaccine'*. Cows have long been vaccinated this way. She points out that masks have to cover the nose and the mouth where it is claimed the 'virus' exists in saliva. Why then don't they take saliva from the mouth as they do with a DNA test instead of pushing a long swab up the nose towards the brain? The ethmoid bone separates the nasal cavity from the brain and within that bone is the cribriform plate. Dr Day says that when the swab is pushed up against this plate and twisted the procedure is 'depositing things back there'. She claims that among these 'things' are nanoparticles that can enter the brain. Researchers have noted that a team at the Gates-funded Johns Hopkins have designed tiny, star-shaped micro-devices that can latch onto intestinal mucosa and release drugs into the body. Mucosa is the thin skin that covers the inside surface of parts of the body such as *the nose* and mouth and produces mucus to protect them. The Johns Hopkins micro-devices are called 'theragrippers' and were 'inspired' by a parasitic worm that digs its sharp teeth into a host's intestines. Nasal swabs are also coated in the sterilisation agent ethylene oxide. The US National Cancer Institute posts this explanation on its website:

> At room temperature, ethylene oxide is a flammable colorless gas with a sweet odor. It is used primarily to produce other chemicals, including antifreeze. In smaller amounts, ethylene oxide is used as a pesticide and a sterilizing agent. The ability of ethylene oxide to damage DNA makes it an effective sterilizing agent but also accounts for its cancer-causing activity.

The Institute mentions lymphoma and leukaemia as cancers most frequently reported to be associated with occupational exposure to ethylene oxide along with stomach and breast cancers. How does anyone think this is going to work out with the constant testing

regime being inflicted on adults and children at home and at school that will accumulate in the body anything that's on the swab?

Doctors know best

It is vital for people to realise that 'hero' doctors 'know' only what the Big Pharma-dominated medical authorities tell them to 'know' and if they refuse to 'know' what they are told to 'know' they are out the door. They are mostly not physicians or healers, but repeaters of the official narrative – or else. I have seen alleged professional doctors on British television make shocking statements that we are supposed to take seriously. One called 'Dr' Amir Khan, who is actually telling patients how to respond to illness, said that men could take the birth pill to 'help slow down the effects of Covid-19'. In March, 2021, another ridiculous 'Covid study' by an American doctor proposed injecting men with the female sex hormone progesterone as a 'Covid' treatment. British doctor Nighat Arif told the BBC that face coverings were now going to be part of ongoing normal. Yes, the vaccine protects you, she said (evidence?) … but the way to deal with viruses in the community was always going to come down to hand washing, face covering and keeping a physical distance. That's not what we were told before the 'vaccine' was circulating. Arif said she couldn't imagine ever again going on the underground or in a lift without a mask. I was just thanking my good luck that she was not my doctor when she said – in March, 2021 – that if 'we are *behaving* and we are doing all the right things' she thought we could 'have our nearest and dearest around us at home … around *Christmas* and N*ew Year!* Her patronising delivery was the usual school teacher talking to six-year-olds as she repeated every government talking point and probably believed them all. If we have learned anything from the 'Covid' experience surely it must be that humanity's perception of doctors needs a fundamental rethink. NHS 'doctor' Sara Kayat told her television audience that the 'Covid vaccine' would '100 percent prevent hospitalisation and death'. Not even Big Pharma claimed that. We have to stop taking 'experts' at their word without question when so many of them are

clueless and only repeating the party line on which their careers depend. That is not to say there are not brilliants doctors – there are and I have spoken to many of them since all this began – but you won't see them in the mainstream media or quoted by the psychopaths and yes-people in government.

Remember the name – Christian Drosten

German virologist Christian Drosten, Director of Charité Institute of Virology in Berlin, became a national star after the pandemic hoax began. He was feted on television and advised the German government on 'Covid' policy. Most importantly to the wider world Drosten led a group that produced the 'Covid' testing protocol for the PCR test. What a remarkable feat given the PCR cannot test for infectious disease and even more so when you think that Drosten said that his method of testing for SARS-CoV-2 was developed 'without having virus material available'. *He developed a test for a 'virus' that he didn't have and had never seen*. Let that sink in as you survey the global devastation that came from what he did. The whole catastrophe of Drosten's 'test' was based on the alleged genetic sequence published by Chinese scientists on the Internet. We will see in the next chapter that this alleged 'genetic sequence' has never been produced by China or anyone and cannot be when there *is no* SARS-CoV-2. Drosten, however, doesn't seem to let little details like that get in the way. He was the lead author with Victor Corman from the same Charité Hospital of the paper 'Detection of 2019 novel coronavirus (2019-nCoV) by real-time PCR' published in a magazine called *Eurosurveillance*. This became known as the Corman-Drosten paper. In November, 2020, with human society devastated by the effects of the Corman-Drosten test baloney, the protocol was publicly challenged by 22 international scientists and independent researchers from Europe, the United States, and Japan. Among them were senior molecular geneticists, biochemists, immunologists, and microbiologists. They produced a document headed 'External peer review of the RTPCR test to detect SARS-Cov-2 Reveals 10 Major Flaws At The Molecular and Methodological Level: Consequences

For False-Positive Results'. The flaws in the Corman-Drosten test included the following:

- The test is non-specific because of erroneous design
- Results are enormously variable
- The test is unable to discriminate between the whole 'virus' and viral fragments
- It doesn't have positive or negative controls
- The test lacks a standard operating procedure
- It is unsupported by proper peer view

The scientists said the PCR 'Covid' testing protocol was not founded on science and they demanded the Corman-Drosten paper be retracted by *Eurosurveillance*. They said all present and previous Covid deaths, cases, and 'infection rates' should be subject to a massive retroactive inquiry. Lockdowns and travel restrictions should be reviewed and relaxed and those diagnosed through PCR to have 'Covid-19' should not be forced to isolate. Dr Kevin Corbett, a health researcher and nurse educator with a long academic career producing a stream of peer-reviewed publications at many UK universities, made the same point about the PCR test debacle. He said of the scientists' conclusions: 'Every scientific rationale for the development of that test has been totally destroyed by this paper. It's like Hiroshima/Nagasaki to the Covid test.' He said that China hadn't given them an isolated 'virus' when Drosten developed the test. Instead they had developed the test from *a sequence in a gene bank*.' Put another way … *they made it up*! The scientists were supported in this contention by a Portuguese appeals court which ruled in November, 2020, that PCR tests are unreliable and it is unlawful to quarantine people based solely on a PCR test. The point about China not providing an isolated virus must be true when the 'virus' has never been isolated to this day and the consequences of that will become clear. Drosten and company produced this useless 'protocol' right on cue in January, 2020, just as the 'virus' was said to

be moving westward and it somehow managed to successfully pass a peer-review in 24 hours. In other words there was no peer-review for a test that would be used to decide who had 'Covid' and who didn't across the world. The Cult-created, Gates-controlled World Health Organization immediately recommended all its nearly 200 member countries to use the Drosten PCR protocol to detect 'cases' and 'deaths'. The sting was underway and it continues to this day.

So who is this Christian Drosten that produced the means through which death, destruction and economic catastrophe would be justified? His education background, including his doctoral thesis, would appear to be somewhat shrouded in mystery and his track record is dire as with another essential player in the 'Covid' hoax, the Gates-funded Professor Neil Ferguson at the Gates-funded Imperial College in London of whom more shortly. Drosten predicted in 2003 that the alleged original SARS 'virus' (SARS-1') was an epidemic that could have serious effects on economies and an effective vaccine would take at least two years to produce. Drosten's answer to every alleged 'outbreak' is a vaccine which you won't be shocked to know. What followed were just 774 official deaths worldwide and none in Germany where there were only nine cases. That is even if you believe there ever was a SARS 'virus' when the evidence is zilch and I will expand on this in the next chapter. Drosten claims to be co-discoverer of 'SARS-1' and developed a test for it in 2003. He was screaming warnings about 'swine flu' in 2009 and how it was a widespread infection far more severe than any dangers from a vaccine could be and people should get vaccinated. It would be helpful for Drosten's vocal chords if he simply recorded the words 'the virus is deadly and you need to get vaccinated' and copies could be handed out whenever the latest made-up threat comes along. Drosten's swine flu epidemic never happened, but Big Pharma didn't mind with governments spending hundreds of millions on vaccines that hardly anyone bothered to use and many who did wished they hadn't. A study in 2010 revealed that the risk of dying from swine flu, or H1N1, was no higher than that of the annual seasonal flu which is what at least most of 'it' really was as in

the case of 'Covid-19'. A media investigation into Drosten asked how with such a record of inaccuracy he could be *the* government adviser on these issues. The answer to that question is the same with Drosten, Ferguson and Fauci – they keep on giving the authorities the 'conclusions' and 'advice' they want to hear. Drosten certainly produced the goods for them in January, 2020, with his PCR protocol garbage and provided the foundation of what German internal medicine specialist Dr Claus Köhnlein, co-author of *Virus Mania*, called the 'test pandemic'. The 22 scientists in the *Eurosurveillance* challenge called out conflicts of interest within the Drosten 'protocol' group and with good reason. Olfert Landt, a regular co-author of Drosten 'studies', owns the biotech company TIB Molbiol Syntheselabor GmbH in Berlin which manufactures and sells the tests that Drosten and his mates come up with. They have done this with SARS, Enterotoxigenic E. coli (ETEC), MERS, Zika 'virus', yellow fever, and now 'Covid'. Landt told the *Berliner Zeitung* newspaper:

> The testing, design and development came from the Charité [Drosten and Corman]. We simply implemented it immediately in the form of a kit. And if we don't have the virus, which originally only existed in Wuhan, we can make a synthetic gene to simulate the genome of the virus. That's what we did very quickly.

This is more confirmation that the Drosten test was designed without access to the 'virus' and only a synthetic simulation which is what SARS-CoV-2 really is – a computer-generated synthetic fiction. It's quite an enterprise they have going here. A Drosten team decides what the test for something should be and Landt's biotech company flogs it to governments and medical systems across the world. His company must have made an absolute fortune since the 'Covid' hoax began. Dr Reiner Fuellmich, a prominent German consumer protection trial lawyer in Germany and California, is on Drosten's case and that of Tedros at the World Health Organization for crimes against humanity with a class-action lawsuit being prepared in the United States and other legal action in Germany.

Why China?

Scamming the world with a 'virus' that doesn't exist would seem impossible on the face of it, but not if you have control of the relatively few people that make policy decisions and the great majority of the global media. Remember it's not about changing 'real' reality it's about controlling *perception* of reality. You don't have to make something happen you only have make people *believe* that it's happening. Renegade Minds understand this and are therefore much harder to swindle. 'Covid-19' is not a 'real' 'virus'. It's a mind virus, like a computer virus, which has infected the minds, not the bodies, of billions. It all started, publically at least, in China and that alone is of central significance. The Cult was behind the revolution led by its asset Mao Zedong, or Chairman Mao, which established the People's Republic of China on October 1st, 1949. It should have been called The Cult's Republic of China, but the name had to reflect the recurring illusion that vicious dictatorships are run by and for the people (see all the 'Democratic Republics' controlled by tyrants). In the same way we have the 'Biden' Democratic Republic of America officially ruled by a puppet tyrant (at least temporarily) on behalf of Cult tyrants. The creation of Mao's merciless communist/fascist dictatorship was part of a frenzy of activity by the Cult at the conclusion of World War Two which, like the First World War, it had instigated through its assets in Germany, Britain, France, the United States and elsewhere. Israel was formed in 1948; the Soviet Union expanded its 'Iron Curtain' control, influence and military power with the Warsaw Pact communist alliance in 1955; the United Nations was formed in 1945 as a Cult precursor to world government; and a long list of world bodies would be established including the World Health Organization (1948), World Trade Organization (1948 under another name until 1995), International Monetary Fund (1945) and World Bank (1944). Human society was redrawn and hugely centralised in the global Problem-Reaction-Solution that was World War Two. All these changes were significant. Israel would become the headquarters of the Sabbatians

and the revolution in China would prepare the ground and control system for the events of 2019/2020.

Renegade Minds know there are no borders except for public consumption. The Cult is a seamless, borderless global entity and to understand the game we need to put aside labels like borders, nations, countries, communism, fascism and democracy. These delude the population into believing that countries are ruled within their borders by a government of whatever shade when these are mere agencies of a global power. America's illusion of democracy and China's communism/fascism are subsidiaries – vehicles – for the same agenda. We may hear about conflict and competition between America and China and on the lower levels that will be true; but at the Cult level they are branches of the same company in the way of the McDonald's example I gave earlier. I have tracked in the books over the years support by US governments of both parties for Chinese Communist Party infiltration of American society through allowing the sale of land, even military facilities, and the acquisition of American business and university influence. All this is underpinned by the infamous stealing of intellectual property and technological know-how. Cult-owned Silicon Valley corporations waive their fraudulent 'morality' to do business with human-rights-free China; Cult-controlled Disney has become China's PR department; and China in effect owns 'American' sports such as basketball which depends for much of its income on Chinese audiences. As a result any sports player, coach or official speaking out against China's horrific human rights record is immediately condemned or fired by the China-worshipping National Basketball Association. One of the first acts of China-controlled Biden was to issue an executive order telling federal agencies to stop making references to the 'virus' by the 'geographic location of its origin'. Long-time Congressman Jerry Nadler warned that criticising China, America's biggest rival, leads to hate crimes against Asian people in the United States. So shut up you bigot. China is fast closing in on Israel as a country that must not be criticised which is apt, really, given that Sabbatians control them both. The two countries have

developed close economic, military, technological and strategic ties which include involvement in China's 'Silk Road' transport and economic initiative to connect China with Europe. Israel was the first country in the Middle East to recognise the establishment of Mao's tyranny in 1950 months after it was established.

Project Wuhan – the 'Covid' Psyop

I emphasise again that the Cult plays the long game and what is happening to the world today is the result of centuries of calculated manipulation following a script to take control step-by-step of every aspect of human society. I will discuss later the common force behind all this that has spanned those centuries and thousands of years if the truth be told. Instigating the Mao revolution in China in 1949 with a 2020 'pandemic' in mind is not only how they work – the 71 years between them is really quite short by the Cult's standards of manipulation preparation. The reason for the Cult's Chinese revolution was to create a fiercely-controlled environment within which an extreme structure for human control could be incubated to eventually be unleashed across the world. We have seen this happen since the 'pandemic' emerged from China with the Chinese control-structure founded on AI technology and tyrannical enforcement sweep across the West. Until the moment when the Cult went for broke in the West and put its fascism on public display Western governments had to pay some lip-service to freedom and democracy to not alert too many people to the tyranny-in-the-making. Freedoms were more subtly eroded and power centralised with covert government structures put in place waiting for the arrival of 2020 when that smokescreen of 'freedom' could be dispensed with. The West was not able to move towards tyranny before 2020 anything like as fast as China which was created as a tyranny and had no limits on how fast it could construct the Cult's blueprint for global control. When the time came to impose that structure on the world it was the same Cult-owned Chinese communist/fascist government that provided the excuse – the 'Covid pandemic'. It was absolutely crucial to the Cult plan for the Chinese response to the 'pandemic' –

draconian lockdowns of the entire population – to become the blueprint that Western countries would follow to destroy the livelihoods and freedom of their people. This is why the Cult-owned, Gates-owned, WHO Director-General Tedros said early on:

The Chinese government is to be congratulated for the extraordinary measures it has taken to contain the outbreak. China is actually setting a new standard for outbreak response and it is not an exaggeration.

Forbes magazine said of China: '… those measures protected untold millions from getting the disease'. The Rockefeller Foundation 'epidemic scenario' document in 2010 said 'prophetically':

However, a few countries did fare better – China in particular. The Chinese government's quick imposition and enforcement of mandatory quarantine for all citizens, as well as its instant and near-hermetic sealing off of all borders, saved millions of lives, stopping the spread of the virus far earlier than in other countries and enabling a swifter post-pandemic recovery.

Once again – *spooky*.

The first official story was the 'bat theory' or rather the bat diversion. The source of the 'virus outbreak' we were told was a ''wet market' in Wuhan where bats and other animals are bought and eaten in horrifically unhygienic conditions. Then another story emerged through the alternative media that the 'virus' had been released on purpose or by accident from a BSL-4 (biosafety level 4) laboratory in Wuhan not far from the wet market. The lab was reported to create and work with lethal concoctions and bioweapons. Biosafety level 4 is the highest in the World Health Organization system of safety and containment. Renegade Minds are aware of what I call designer manipulation. The ideal for the Cult is for people to buy its prime narrative which in the opening salvoes of the 'pandemic' was the wet market story. It knows, however, that there is now a considerable worldwide alternative media of researchers sceptical of anything governments say and they are often given a version of events in a form they can perceive as credible while misdirecting them from the real truth. In this case let them

think that the conspiracy involved is a 'bioweapon virus' released from the Wuhan lab to keep them from the real conspiracy – *there is no 'virus'*. The WHO's current position on the source of the outbreak at the time of writing appears to be: 'We haven't got a clue, mate.' This is a good position to maintain mystery and bewilderment. The inner circle will know where the 'virus' came from – *nowhere*. The bottom line was to ensure the public believed there *was* a 'virus' and it didn't much matter if they thought it was natural or had been released from a lab. The belief that there was a 'deadly virus' was all that was needed to trigger global panic and fear. The population was terrified into handing their power to authority and doing what they were told. They had to or they were 'all gonna die'.

In March, 2020, information began to come my way from real doctors and scientists and my own additional research which had my intuition screaming: 'Yes, that's it! *There is no virus*.' The 'bioweapon' was not the 'virus'; it was the *'vaccine'* already being talked about that would be the bioweapon. My conclusion was further enhanced by happenings in Wuhan. The 'virus' was said to be sweeping the city and news footage circulated of people collapsing in the street (which they've never done in the West with the same 'virus'). The Chinese government was building 'new hospitals' in a matter of ten days to 'cope with demand' such was the virulent nature of the 'virus'. Yet in what seemed like no time the 'new hospitals' closed – even if they even opened – and China declared itself 'virus-free'. It was back to business as usual. This was more propaganda to promote the Chinese draconian lockdowns in the West as the way to 'beat the virus'. Trouble was that we subsequently had lockdown after lockdown, but never business as usual. As the people of the West and most of the rest of the world were caught in an ever-worsening spiral of lockdown, social distancing, masks, isolated old people, families forced apart, and livelihood destruction, it was party-time in Wuhan. Pictures emerged of thousands of people enjoying pool parties and concerts. It made no sense until you realised there never was a 'virus' and the

whole thing was a Cult set-up to transform human society out of one its major global strongholds – China.

How is it possible to deceive virtually the entire world population into believing there is a deadly virus when there is not even a 'virus' let alone a deadly one? It's nothing like as difficult as you would think and that's clearly true because it happened.

Postscript: See end of book Postscript for more on the 'Wuhan lab virus release' story which the authorities and media were pushing heavily in the summer of 2021 to divert attention from the truth that the 'Covid virus' is pure invention.

CHAPTER FIVE

There *is no* 'virus'

You can fool some of the people all of the time, and all of the people some of the time, but you cannot fool all of the people all of the time
Abraham Lincoln

The greatest form of mind control is repetition. The more you repeat the same mantra of alleged 'facts' the more will accept them to be true. It becomes an 'everyone knows that, mate'. If you can also censor any other version or alternative to your alleged 'facts' you are pretty much home and cooking.

By the start of 2020 the Cult owned the global mainstream media almost in its entirety to spew out its 'Covid' propaganda and ignore or discredit any other information and view. Cult-owned social media platforms in Cult-owned Silicon Valley were poised and ready to unleash a campaign of ferocious censorship to obliterate all but the official narrative. To complete the circle many demands for censorship by Silicon Valley were led by the mainstream media as 'journalists' became full-out enforcers for the Cult both as propagandists and censors. Part of this has been the influx of young people straight out of university who have become 'journalists' in significant positions. They have no experience and a headful of programmed perceptions from their years at school and university at a time when today's young are the most perceptually-targeted generations in known human history given the insidious impact of technology. They enter the media perceptually prepared and ready to repeat the narratives of the system that programmed them to

repeat its narratives. The BBC has a truly pathetic 'specialist disinformation reporter' called Marianna Spring who fits this bill perfectly. She is clueless about the world, how it works and what is really going on. Her role is to discredit anyone doing the job that a proper journalist would do and system-serving hacks like Spring wouldn't dare to do or even see the need to do. They are too busy licking the arse of authority which can never be wrong and, in the case of the BBC propaganda programme, *Panorama*, contacting payments systems such as PayPal to have a donations page taken down for a film company making documentaries questioning vaccines. Even the BBC soap opera *EastEnders* included a disgracefully biased scene in which an inarticulate white working class woman was made to look foolish for questioning the 'vaccine' while a well-spoken black man and Asian woman promoted the government narrative. It ticked every BBC box and the fact that the black and minority community was resisting the 'vaccine' had nothing to do with the way the scene was written. The BBC has become a disgusting tyrannical propaganda and censorship operation that should be defunded and disbanded and a free media take its place with a brief to stop censorship instead of demanding it. A BBC 'interview' with Gates goes something like: 'Mr Gates, sir, if I can call you sir, would you like to tell our audience why you are such a great man, a wonderful humanitarian philanthropist, and why you should absolutely be allowed as a software salesman to decide health policy for approaching eight billion people? Thank you, sir, please sir.' Propaganda programming has been incessant and merciless and when all you hear is the same story from the media, repeated by those around you who have only heard the same story, is it any wonder that people on a grand scale believe absolute mendacious garbage to be true? You are about to see, too, why this level of information control is necessary when the official 'Covid' narrative is so nonsensical and unsupportable by the evidence.

Structure of Deceit

The pyramid structure through which the 'Covid' hoax has been manifested is very simple and has to be to work. As few people as possible have to be involved with full knowledge of what they are doing – and why – or the real story would get out. At the top of the pyramid are the inner core of the Cult which controls Bill Gates who, in turn, controls the World Health Organization through his pivotal funding and his puppet Director-General mouthpiece, Tedros. Before he was appointed Tedros was chair of the Gates-founded Global Fund to 'fight against AIDS, tuberculosis and malaria', a board member of the Gates-funded 'vaccine alliance' GAVI, and on the board of another Gates-funded organisation. Gates owns him and picked him for a specific reason – Tedros is a crook and worse. 'Dr' Tedros (he's not a medical doctor, the first WHO chief not to be) was a member of the tyrannical Marxist government of Ethiopia for decades with all its human rights abuses. He has faced allegations of corruption and misappropriation of funds and was exposed three times for covering up cholera epidemics while Ethiopia's health minister. Tedros appointed the mass-murdering genocidal Zimbabwe dictator Robert Mugabe as a WHO goodwill ambassador for public health which, as with Tedros, is like appointing a psychopath to run a peace and love campaign. The move was so ridiculous that he had to drop Mugabe in the face of widespread condemnation. American economist David Steinman, a Nobel peace prize nominee, lodged a complaint with the International Criminal Court in The Hague over alleged genocide by Tedros when he was Ethiopia's foreign minister. Steinman says Tedros was a 'crucial decision maker' who directed the actions of Ethiopia's security forces from 2013 to 2015 and one of three officials in charge when those security services embarked on the 'killing' and 'torturing' of Ethiopians. You can see where Tedros is coming from and it's sobering to think that he has been the vehicle for Gates and the Cult to direct the global response to 'Covid'. Think about that. A psychopathic Cult dictates to psychopath Gates who dictates to psychopath Tedros who dictates how countries of the world must respond to a 'Covid virus' never scientifically shown to exist. At the same time psychopathic Cult-owned Silicon Valley information

giants like Google, YouTube, Facebook and Twitter announced very early on that they would give the Cult/Gates/Tedros/WHO version of the narrative free advertising and censor those who challenged their intelligence-insulting, mendacious story.

The next layer in the global 'medical' structure below the Cult, Gates and Tedros are the chief medical officers and science 'advisers' in each of the WHO member countries which means virtually all of them. Medical officers and arbiters of science (they're not) then take the WHO policy and recommended responses and impose them on their country's population while the political 'leaders' say they are deciding policy (they're clearly not) by 'following the science' on the advice of the 'experts' – the same medical officers and science 'advisers' (dictators). In this way with the rarest of exceptions the entire world followed the same policy of lockdown, people distancing, masks and 'vaccines' dictated by the psychopathic Cult, psychopathic Gates and psychopathic Tedros who we are supposed to believe give a damn about the health of the world population they are seeking to enslave. That, amazingly, is all there is to it in terms of crucial decision-making. Medical staff in each country then follow like sheep the dictates of the shepherds at the top of the national medical hierarchies – chief medical officers and science 'advisers' who themselves follow like sheep the shepherds of the World Health Organization and the Cult. Shepherds at the national level often have major funding and other connections to Gates and his Bill and Melinda Gates Foundation which carefully hands out money like confetti at a wedding to control the entire global medical system from the WHO down.

Follow the money

Christopher Whitty, Chief Medical Adviser to the UK Government at the centre of 'virus' policy, a senior adviser to the government's Scientific Advisory Group for Emergencies (SAGE), and Executive Board member of the World Health Organization, was gifted a grant of $40 million by the Bill and Melinda Gates Foundation for malaria research in Africa. The BBC described the unelected Whitty as 'the

official who will probably have the greatest impact on our everyday lives of any individual policymaker in modern times' and so it turned out. What Gates and Tedros have said Whitty has done like his equivalents around the world. Patrick Vallance, co-chair of SAGE and the government's Chief Scientific Adviser, is a former executive of Big Pharma giant GlaxoSmithKline with its fundamental financial and business connections to Bill Gates. In September, 2020, it was revealed that Vallance owned a deferred bonus of shares in GlaxoSmithKline worth £600,000 while the company was 'developing' a 'Covid vaccine'. Move along now – nothing to see here – what could possibly be wrong with that? Imperial College in London, a major player in 'Covid' policy in Britain and elsewhere with its 'Covid-19' Response Team, is funded by Gates and has big connections to China while the now infamous Professor Neil Ferguson, the useless 'computer modeller' at Imperial College is also funded by Gates. Ferguson delivered the dramatically inaccurate excuse for the first lockdowns (much more in the next chapter). The Institute for Health Metrics and Evaluation (IHME) in the United States, another source of outrageously false 'Covid' computer models to justify lockdowns, is bankrolled by Gates who is a vehement promotor of lockdowns. America's version of Whitty and Vallance, the again now infamous Anthony Fauci, has connections to 'Covid vaccine' maker Moderna as does Bill Gates through funding from the Bill and Melinda Gates Foundation. Fauci is director of the National Institute of Allergy and Infectious Diseases (NIAID), a major recipient of Gates money, and they are very close. Deborah Birx who was appointed White House Coronavirus Response Coordinator in February, 2020, is yet another with ties to Gates. Everywhere you look at the different elements around the world behind the coordination and decision making of the 'Covid' hoax there is Bill Gates and his money. They include the World Health Organization; Centers for Disease Control (CDC) in the United States; National Institutes of Health (NIH) of Anthony Fauci; Imperial College and Neil Ferguson; the London School of Hygiene where Chris Whitty worked; Regulatory agencies like the UK Medicines & Healthcare products Regulatory Agency (MHRA)

which gave emergency approval for 'Covid vaccines'; Wellcome Trust; GAVI, the Vaccine Alliance; the Coalition for Epidemic Preparedness Innovations (CEPI); Johns Hopkins University which has compiled the false 'Covid' figures; and the World Economic Forum. A Nationalfile.com article said:

> Gates has a lot of pull in the medical world, he has a multi-million dollar relationship with Dr. Fauci, and Fauci originally took the Gates line supporting vaccines and casting doubt on [the drug hydroxychloroquine]. Coronavirus response team member Dr. Deborah Birx, appointed by former president Obama to serve as United States Global AIDS Coordinator, also sits on the board of a group that has received billions from Gates' foundation, and Birx reportedly used a disputed Bill Gates-funded model for the White House's Coronavirus effort. Gates is a big proponent for a population lockdown scenario for the Coronavirus outbreak.

Another funder of Moderna is the Defense Advanced Research Projects Agency (DARPA), the technology-development arm of the Pentagon and one of the most sinister organisations on earth. DARPA had a major role with the CIA covert technology-funding operation In-Q-Tel in the development of Google and social media which is now at the centre of global censorship. Fauci and Gates are extremely close and openly admit to talking regularly about 'Covid' policy, but then why wouldn't Gates have a seat at every national 'Covid' table after his Foundation committed $1.75 billion to the 'fight against Covid-19'. When passed through our Orwellian Translation Unit this means that he has bought and paid for the Cult-driven 'Covid' response worldwide. Research the major 'Covid' response personnel in your own country and you will find the same Gates funding and other connections again and again. Medical and science chiefs following World Health Organization 'policy' sit atop a medical hierarchy in their country of administrators, doctors and nursing staff. These 'subordinates' are told they must work and behave in accordance with the policy delivered from the 'top' of the national 'health' pyramid which is largely the policy delivered by the WHO which is the policy delivered by Gates and the Cult. The whole 'Covid' narrative has been imposed on medical staff by a climate of fear although great numbers don't even need that to comply. They do so through breathtaking levels of ignorance and

include doctors who go through life simply repeating what Big Pharma and their hierarchical masters tell them to say and believe. No wonder Big Pharma 'medicine' is one of the biggest killers on Planet Earth.

The same top-down system of intimidation operates with regard to the Cult Big Pharma cartel which also dictates policy through national and global medical systems in this way. The Cult and Big Pharma agendas are the same because the former controls and owns the latter. 'Health' administrators, doctors, and nursing staff are told to support and parrot the dictated policy or they will face consequences which can include being fired. How sad it's been to see medical staff meekly repeating and imposing Cult policy without question and most of those who can see through the deceit are only willing to speak anonymously off the record. They know what will happen if their identity is known. This has left the courageous few to expose the lies about the 'virus', face masks, overwhelmed hospitals that aren't, and the dangers of the 'vaccine' that isn't a vaccine. When these medical professionals and scientists, some renowned in their field, have taken to the Internet to expose the truth their articles, comments and videos have been deleted by Cult-owned Facebook, Twitter and YouTube. What a real head-shaker to see YouTube videos with leading world scientists and highly qualified medical specialists with an added link underneath to the notorious Cult propaganda website *Wikipedia* to find the 'facts' about the same subject.

HIV – the 'Covid' trial-run

I'll give you an example of the consequences for health and truth that come from censorship and unquestioning belief in official narratives. The story was told by PCR inventor Kary Mullis in his book *Dancing Naked in the Mind Field*. He said that in 1984 he accepted as just another scientific fact that Luc Montagnier of France's Pasteur Institute and Robert Gallo of America's National Institutes of Health had independently discovered that a 'retrovirus' dubbed HIV (human immunodeficiency virus) caused AIDS. They

were, after all, Mullis writes, specialists in retroviruses. This is how the medical and science pyramids work. Something is announced or *assumed* and then becomes an everybody-knows-that purely through repetition of the assumption as if it is fact. Complete crap becomes accepted truth with no supporting evidence and only repetition of the crap. This is how a 'virus' that doesn't exist became the 'virus' that changed the world. The HIV-AIDS fairy story became a multi-billion pound industry and the media poured out propaganda terrifying the world about the deadly HIV 'virus' that caused the lethal AIDS. By then Mullis was working at a lab in Santa Monica, California, to detect retroviruses with his PCR test in blood donations received by the Red Cross. In doing so he asked a virologist where he could find a reference for HIV being the cause of AIDS. 'You don't need a reference,' the virologist said … *'Everybody knows it.'* Mullis said he wanted to quote a reference in the report he was doing and he said he felt a little funny about not knowing the source of such an important discovery when everyone else seemed to. The virologist suggested he cite a report by the Centers for Disease Control and Prevention (CDC) on morbidity and mortality. Mullis read the report, but it only said that an organism had been identified and did not say how. The report did not identify the original scientific work. Physicians, however, *assumed* (key recurring theme) that if the CDC was convinced that HIV caused AIDS then proof must exist. Mullis continues:

I did computer searches. Neither Montagnier, Gallo, nor anyone else had published papers describing experiments which led to the conclusion that HIV probably caused AIDS. I read the papers in Science for which they had become well known as AIDS doctors, but all they had said there was that they had found evidence of a past infection by something which was probably HIV in some AIDS patients.

They found antibodies. Antibodies to viruses had always been considered evidence of past disease, not present disease. Antibodies signaled that the virus had been defeated. The patient had saved himself. There was no indication in these papers that this virus caused a disease. They didn't show that everybody with the antibodies had the disease. In fact they found some healthy people with antibodies.

Mullis asked why their work had been published if Montagnier and Gallo hadn't really found this evidence, and why had they been fighting so hard to get credit for the discovery? He says he was hesitant to write 'HIV is the probable cause of AIDS' until he found published evidence to support that. 'Tens of thousands of scientists and researchers were spending billions of dollars a year doing research based on this idea,' Mullis writes. 'The reason had to be there somewhere; otherwise these people would not have allowed their research to settle into one narrow channel of investigation.' He said he lectured about PCR at numerous meetings where people were always talking about HIV and he asked them how they knew that HIV was the cause of AIDS:

> Everyone said something. Everyone had the answer at home, in the office, in some drawer. They all knew, and they would send me the papers as soon as they got back. But I never got any papers. Nobody ever sent me the news about how AIDS was caused by HIV.

Eventually Mullis was able to ask Montagnier himself about the reference proof when he lectured in San Diego at the grand opening of the University of California AIDS Research Center. Mullis says this was the last time he would ask his question without showing anger. Montagnier said he should reference the CDC report. 'I read it', Mullis said, and it didn't answer the question. 'If Montagnier didn't know the answer who the hell did?' Then one night Mullis was driving when an interview came on National Public Radio with Peter Duesberg, a prominent virologist at Berkeley and a California Scientist of the Year. Mullis says he finally understood why he could not find references that connected HIV to AIDS – *there weren't any!* No one had ever proved that HIV causes AIDS even though it had spawned a multi-billion pound global industry and the media was repeating this as fact every day in their articles and broadcasts terrifying the shit out of people about AIDS and giving the impression that a positive test for HIV (see 'Covid') was a death sentence. Duesberg was a threat to the AIDS gravy train and the agenda that underpinned it. He was therefore abused and castigated after he told the Proceedings of the National Academy of Sciences

there was no good evidence implicating the new 'virus'. Editors rejected his manuscripts and his research funds were deleted. Mullis points out that the CDC has defined AIDS as one of more than 30 diseases *if accompanied* by a positive result on a test that detects antibodies to HIV; but those same diseases are not defined as AIDS cases when antibodies are not detected:

> If an HIV-positive woman develops uterine cancer, for example, she is considered to have AIDS. If she is not HIV positive, she simply has uterine cancer. An HIV-positive man with tuberculosis has AIDS; if he tests negative he simply has tuberculosis. If he lives in Kenya or Colombia, where the test for HIV antibodies is too expensive, he is simply presumed to have the antibodies and therefore AIDS, and therefore he can be treated in the World Health Organization's clinic. It's the only medical help available in some places. And it's free, because the countries that support WHO are worried about AIDS.

Mullis accuses the CDC of continually adding new diseases (see ever more 'Covid symptoms') to the grand AIDS definition and of virtually doctoring the books to make it appear as if the disease continued to spread. He cites how in 1993 the CDC enormously broadened its AIDS definition and county health authorities were delighted because they received $2,500 per year from the Federal government for every reported AIDS case. Ladies and gentlemen, I have just described, via Kary Mullis, the 'Covid pandemic' of 2020 and beyond. Every element is the same and it's been pulled off in the same way by the same networks.

The 'Covid virus' exists? Okay – prove it. Er … still waiting

What Kary Mullis described with regard to 'HIV' has been repeated with 'Covid'. A claim is made that a new, or 'novel', infection has been found and the entire medical system of the world repeats that as fact exactly as they did with HIV and AIDS. No one in the mainstream asks rather relevant questions such as 'How do you know?' and 'Where is your proof?' The SARS-Cov-2 'virus' and the 'Covid-19 disease' became an overnight 'everybody-knows-that'. The origin could be debated and mulled over, but what you could not suggest was that 'SARS-Cov-2' didn't exist. That would be

ridiculous. 'Everybody knows' the 'virus' exists. Well, I didn't for one along with American proper doctors like Andrew Kaufman and Tom Cowan and long-time American proper journalist Jon Rappaport. We dared to pursue the obvious and simple question: 'Where's the evidence?' The overwhelming majority in medicine, journalism and the general public did not think to ask that. After all, *everyone knew* there was a new 'virus'. Everyone was saying so and I heard it on the BBC. Some would eventually argue that the 'deadly virus' was nothing like as deadly as claimed, but few would venture into the realms of its very existence. Had they done so they would have found that the evidence for that claim had gone AWOL as with HIV causes AIDS. In fact, not even that. For something to go AWOL it has to exist in the first place and scientific proof for a 'SARS-Cov-2' can be filed under nothing, nowhere and zilch.

Dr Andrew Kaufman is a board-certified forensic psychiatrist in New York State, a Doctor of Medicine and former Assistant Professor and Medical Director of Psychiatry at SUNY Upstate Medical University, and Medical Instructor of Hematology and Oncology at the Medical School of South Carolina. He also studied biology at the Massachusetts Institute of Technology (MIT) and trained in Psychiatry at Duke University. Kaufman is retired from allopathic medicine, but remains a consultant and educator on natural healing, I saw a video of his very early on in the 'Covid' hoax in which he questioned claims about the 'virus' in the absence of any supporting evidence and with plenty pointing the other way. I did everything I could to circulate his work which I felt was asking the pivotal questions that needed an answer. I can recommend an excellent pull-together interview he did with the website The Last Vagabond entitled *Dr Andrew Kaufman: Virus Isolation, Terrain Theory and Covid-19* and his website is andrewkaufmanmd.com. Kaufman is not only a forensic psychiatrist; he is forensic in all that he does. He always reads original scientific papers, experiments and studies instead of second-third-fourth-hand reports about the 'virus' in the media which are repeating the repeated repetition of the narrative. When he did so with the original Chinese 'virus' papers Kaufman

realised that there was no evidence of a 'SARS-Cov-2'. They had never – from the start – shown it to exist and every repeat of this claim worldwide was based on the accepted existence of proof that was nowhere to be found – see Kary Mullis and HIV. Here we go again.

Let's postulate

Kaufman discovered that the Chinese authorities immediately concluded that the cause of an illness that broke out among about 200 initial patients in Wuhan was a 'new virus' when there were no grounds to make that conclusion. The alleged 'virus' was not isolated from other genetic material in their samples and then shown through a system known as Koch's postulates to be the causative agent of the illness. The world was told that the SARS-Cov-2 'virus' caused a disease they called 'Covid-19' which had 'flu-like' symptoms and could lead to respiratory problems and pneumonia. If it wasn't so tragic it would almost be funny. *'Flu-like' symptoms'? Pneumonia? Respiratory disease?* What in *CHINA* and particularly in *Wuhan*, one of the most polluted cities in the world with a resulting epidemic of respiratory disease?? Three hundred thousand people get pneumonia in China every year and there are nearly a billion cases worldwide of 'flu-like symptoms'. These have a whole range of causes – including pollution in Wuhan – but no other possibility was credibly considered in late 2019 when the world was told there was a new and deadly 'virus'. The global prevalence of pneumonia and 'flu-like systems' gave the Cult networks unlimited potential to re-diagnose these other causes as the mythical 'Covid-19' and that is what they did from the very start. Kaufman revealed how Chinese medical and science authorities (all subordinates to the Cult-owned communist government) took genetic material from the lungs of only a few of the first patients. The material contained their own cells, bacteria, fungi and other microorganisms living in their bodies. The only way you could prove the existence of the 'virus' and its responsibility for the alleged 'Covid-19' was to isolate the virus from all the other material – a process also known as 'purification' – and

then follow the postulates sequence developed in the late 19th century by German physician and bacteriologist Robert Koch which became the 'gold standard' for connecting an alleged causation agent to a disease:

1. The microorganism (bacteria, fungus, virus, etc.) must be present in every case of the disease and all patients must have the same symptoms. It must also *not be present in healthy individuals*.

2. The microorganism must be isolated from the host with the disease. If the microorganism is a bacteria or fungus it must be grown in a pure culture. If it is a virus, it must be purified (i.e. containing no other material except the virus particles) from a clinical sample.

3. The specific disease, with all of its characteristics, must be reproduced when the infectious agent (the purified virus or a pure culture of bacteria or fungi) is inoculated into a healthy, susceptible host.

4. The microorganism must be recoverable from the experimentally infected host as in step 2.

Not one of these criteria has been met in the case of 'SARS-Cov-2' and 'Covid-19'. Not ONE. *EVER*. Robert Koch refers to bacteria and not viruses. What are called 'viral particles' are so minute (hence masks are useless by any definition) that they could only be seen after the invention of the electron microscope in the 1930s and can still only be observed through that means. American bacteriologist and virologist Thomas Milton Rivers, the so-called 'Father of Modern Virology' who was very significantly director of the Rockefeller Institute for Medical Research in the 1930s, developed a less stringent version of Koch's postulates to identify 'virus' causation known as 'Rivers criteria'. 'Covid' did not pass that process either. Some even doubt whether any 'virus' can be isolated from other particles containing genetic material in the Koch method. Freedom of Information requests in many countries asking for scientific proof that the 'Covid virus' has been purified and isolated and shown to exist have all come back with a 'we don't have that' and when this happened with a request to the UK Department of Health they added this comment:

However, outside of the scope of the [Freedom of Information Act] and on a discretionary basis, the following information has been advised to us, which may be of interest. Most infectious diseases are caused by viruses, bacteria or fungi. Some bacteria or fungi have the capacity to grow on their own in isolation, for example in colonies on a petri dish. Viruses are different in that they are what we call 'obligate pathogens' – that is, they cannot survive or reproduce without infecting a host ...

… For some diseases, it is possible to establish causation between a microorganism and a disease by isolating the pathogen from a patient, growing it in pure culture and reintroducing it to a healthy organism. These are known as 'Koch's postulates' and were developed in 1882. However, as our understanding of disease and different disease-causing agents has advanced, these are no longer the method for determining causation [Andrew Kaufman asks why in that case are there two published articles falsely claiming to satisfy Koch's postulates].

It has long been known that viral diseases cannot be identified in this way as viruses cannot be grown in 'pure culture'. When a patient is tested for a viral illness, this is normally done by looking for the presence of antigens, or viral genetic code in a host with molecular biology techniques [Kaufman asks how you could know the origin of these chemicals without having a pure culture for comparison].

For the record 'antigens' are defined so:

Invading microorganisms have antigens on their surface that the human body can recognise as being foreign – meaning not belonging to it. When the body recognises a foreign antigen, lymphocytes (white blood cells) produce antibodies, which are complementary in shape to the antigen.

Notwithstanding that this is open to question in relation to 'SARS-Cov-2' the presence of 'antibodies' can have many causes and they are found in people that are perfectly well. Kary Mullis said: 'Antibodies … had always been considered evidence of past disease, not present disease.'

'Covid' really is a *computer* 'virus'

Where the UK Department of Health statement says 'viruses' are now 'diagnosed' through a 'viral genetic code in a host with molecular biology techniques', they mean … *the PCR test* which its inventor said cannot test for infectious disease. They have no credible method of connecting a 'virus' to a disease and we will see that there is no scientific proof that any 'virus' causes any disease or there is any such thing as a 'virus' in the way that it is described. Tenacious Canadian researcher Christine Massey and her team made

some 40 Freedom of Information requests to national public health agencies in different countries asking for proof that SARS-CoV-2 has been isolated and not one of them could supply that information. Massey said of her request in Canada: 'Freedom of Information reveals Public Health Agency of Canada has no record of 'SARS-COV-2' isolation performed by anyone, anywhere, ever.' If you accept the comment from the UK Department of Health it's because they can't isolate a 'virus'. Even so many 'science' papers claimed to have isolated the 'Covid virus' until they were questioned and had to admit they hadn't. A reply from the Robert Koch Institute in Germany was typical: 'I am not aware of a paper which purified isolated SARS-CoV-2.' So what the hell was Christian Drosten and his gang using to design the 'Covid' testing protocol that has produced all the illusory Covid' cases and 'Covid' deaths when the head of the Chinese version of the CDC admitted there was a problem right from the start in that the 'virus' had never been isolated/purified? Breathe deeply: What they are calling 'Covid' is actually created by a *computer program* i.e. *they made it up* – er, that's it. They took lung fluid, with many sources of genetic material, from one single person alleged to be infected with Covid-19 by a PCR test which they *claimed*, without clear evidence, contained a 'virus'. They used several computer programs to create a model of a theoretical virus genome sequence from more than fifty-six million small sequences of RNA, each of an unknown source, assembling them like a puzzle with no known solution. The computer filled in the gaps with sequences from bits in the gene bank to make it look like a bat SARS-like coronavirus! A wave of the magic wand and poof, an in silico (computer-generated) genome, a scientific fantasy, was created. UK health researcher Dr Kevin Corbett made the same point with this analogy:

> … It's like giving you a few bones and saying that's your fish. It could be any fish. Not even a skeleton. Here's a few fragments of bones. That's your fish … It's all from gene bank and the bits of the virus sequence that weren't there they made up.
>
> They synthetically created them to fill in the blanks. That's what genetics is; it's a code. So it's ABBBCCDDD and you're missing some what you think is EEE so you put it in. It's all

synthetic. You just manufacture the bits that are missing. This is the end result of the geneticization of virology. This is basically a computer virus.

Further confirmation came in an email exchange between British citizen journalist Frances Leader and the government's Medicines & Healthcare Products Regulatory Agency (the Gates-funded MHRA) which gave emergency permission for untested 'Covid vaccines' to be used. The agency admitted that the 'vaccine' is not based on an isolated 'virus', but comes from a *computer-generated model*. Frances Leader was naturally banned from Cult-owned fascist Twitter for making this exchange public. The process of creating computer-generated alleged 'viruses' is called 'in silico' or 'in silicon' – computer chips – and the term 'in silico' is believed to originate with biological experiments using only a computer in 1989. 'Vaccines' involved with 'Covid' are also produced 'in silico' or by computer not a natural process. If the original 'virus' is nothing more than a made-up computer model how can there be 'new variants' of something that never existed in the first place? They are not new 'variants'; they are new *computer models* only minutely different to the original program and designed to further terrify the population into having the 'vaccine' and submitting to fascism. You want a 'new variant'? Click, click, enter – there you go. Tell the medical profession that you have discovered a 'South African variant', 'UK variants' or a 'Brazilian variant' and in the usual HIV-causes-AIDS manner they will unquestioningly repeat it with no evidence whatsoever to support these claims. They will go on television and warn about the dangers of 'new variants' while doing nothing more than repeating what they have been told to be true and knowing that any deviation from that would be career suicide. Big-time insiders will know it's a hoax, but much of the medical community is clueless about the way they are being played and themselves play the public without even being aware they are doing so. What an interesting 'coincidence' that AstraZeneca and Oxford University were conducting 'Covid vaccine trials' in the three countries – the UK, South Africa and Brazil – where the first three 'variants' were claimed to have 'broken out'.

Here's your 'virus' – it's a unicorn

Dr Andrew Kaufman presented a brilliant analysis describing how the 'virus' was imagined into fake existence when he dissected an article published by *Nature* and written by 19 authors detailing *alleged* 'sequencing of a complete viral genome' of the 'new SARS-CoV-2 virus'. This computer-modelled *in silico* genome was used as a template for all subsequent genome sequencing experiments that resulted in the so-called variants which he said now number more than *6,000*. The fake genome was constructed from more than 56 million individual short strands of RNA. Those little pieces were assembled into longer pieces by finding areas of overlapping sequences. The computer programs created over two million possible combinations from which the authors simply chose the longest one. They then compared this to a 'bat virus' and the computer 'alignment' rearranged the sequence and filled in the gaps! They called this computer-generated abomination the 'complete genome'. Dr Tom Cowan, a fellow medical author and collaborator with Kaufman, said such computer-generation constitutes scientific fraud and he makes this superb analogy:

> Here is an equivalency: A group of researchers claim to have found a unicorn because they found a piece of a hoof, a hair from a tail, and a snippet of a horn. They then add that information into a computer and program it to re-create the unicorn, and they then claim this computer re-creation is the real unicorn. Of course, they had never actually seen a unicorn so could not possibly have examined its genetic makeup to compare their samples with the actual unicorn's hair, hooves and horn.
>
> The researchers claim they decided which is the real genome of SARS-CoV-2 by 'consensus', sort of like a vote. Again, different computer programs will come up with different versions of the imaginary 'unicorn', so they come together as a group and decide which is the real imaginary unicorn.

This is how the 'virus' that has transformed the world was brought into fraudulent 'existence'. Extraordinary, yes, but as the Nazis said the bigger the lie the more will believe it. Cowan, however, wasn't finished and he went on to identify what he called the real blockbuster in the paper. He quotes this section from a paper written

by virologists and published by the CDC and then explains what it means:

Therefore, we examined the capacity of SARS-CoV-2 to infect and replicate in several common primate and human cell lines, including human adenocarcinoma cells (A549), human liver cells (HUH 7.0), and human embryonic kidney cells (HEK-293T). In addition to Vero E6 and Vero CCL81 cells. ... Each cell line was inoculated at high multiplicity of infection and examined 24h post-infection.

No CPE was observed in any of the cell lines except in Vero cells, which grew to greater than 10 to the 7th power at 24 h post-infection. In contrast, HUH 7.0 and 293T showed only modest viral replication, and A549 cells were incompatible with SARS CoV-2 infection.

Cowan explains that when virologists attempt to prove infection they have three possible 'hosts' or models on which they can test. The first was humans. Exposure to humans was generally not done for ethical reasons and has never been done with SARS-CoV-2 or any coronavirus. The second possible host was animals. Cowan said that forgetting for a moment that they never actually use purified virus when exposing animals they do use solutions that they *claim* contain the virus. Exposure to animals has been done with SARS-CoV-2 in an experiment involving mice and this is what they found: *None of the wild (normal) mice got sick*. In a group of genetically-modified mice, a statistically insignificant number lost weight and had slightly bristled fur, but they experienced nothing like the illness called 'Covid-19'. Cowan said the third method – the one they mostly rely on – is to inoculate solutions they *say* contain the virus onto a variety of tissue cultures. This process had never been shown to kill tissue *unless* the sample material was starved of nutrients and poisoned as *part of the process*. Yes, incredibly, in tissue experiments designed to show the 'virus' is responsible for killing the tissue they starve the tissue of nutrients and add toxic drugs including antibiotics and they do not have control studies to see if it's the starvation and poisoning that is degrading the tissue rather than the 'virus' they allege to be in there somewhere. You want me to pinch you? Yep, I understand. Tom Cowan said this about the whole nonsensical farce as he explains what that quote from the CDC paper really means:

The shocking thing about the above quote is that using their own methods, the virologists found that solutions containing SARS-CoV-2 – even in high amounts – were NOT, I repeat NOT, infective to any of the three human tissue cultures they tested. In plain English, this means they proved, on their terms, that this 'new coronavirus' is not infectious to human beings. It is ONLY infective to monkey kidney cells, and only then when you add two potent drugs (gentamicin and amphotericin), known to be toxic to kidneys, to the mix.

My friends, read this again and again. These virologists, published by the CDC, performed a clear proof, on their terms, showing that the SARS-CoV-2 virus is harmless to human beings. That is the only possible conclusion, but, unfortunately, this result is not even mentioned in their conclusion. They simply say they can provide virus stocks cultured only on monkey Vero cells, thanks for coming.

Cowan concluded: 'If people really understood how this "science" was done, I would hope they would storm the gates and demand honesty, transparency and truth.' Dr Michael Yeadon, former Vice President and Chief Scientific Adviser at drug giant Pfizer has been a vocal critic of the 'Covid vaccine' and its potential for multiple harm. He said in an interview in April, 2021, that 'not one [vaccine] has the virus. He was asked why vaccines normally using a 'dead' version of a disease to activate the immune system were not used for 'Covid' and instead we had the synthetic methods of the 'mRNA Covid vaccine'. Yeadon said that to do the former 'you'd have to have some of [the virus] wouldn't you?' He added: 'No-one's got any – seriously.' Yeadon said that surely they couldn't have fooled the whole world for a year without having a virus, 'but oddly enough ask around – no one's got it'. He didn't know why with all the 'great labs' around the world that the virus had not been isolated – 'Maybe they've been too busy running bad PCR tests and vaccines that people don't need.' What is today called 'science' is not 'science' at all. Science is no longer what is, but whatever people can be manipulated to *believe* that it is. Real science has been hijacked by the Cult to dispense and produce the 'expert scientists' and contentions that suit the agenda of the Cult. How big-time this has happened with the 'Covid' hoax which is entirely based on fake science delivered by fake 'scientists' and fake 'doctors'. The human-caused climate change hoax is also entirely based on fake science delivered by fake 'scientists' and fake 'climate experts'. In both cases real

scientists, climate experts and doctors have their views suppressed and deleted by the Cult-owned science establishment, media and Silicon Valley. This is the 'science' that politicians claim to be 'following' and a common denominator of 'Covid' and climate are Cult psychopaths Bill Gates and his mate Klaus Schwab at the Gates-funded World Economic Forum. But, don't worry, it's all just a coincidence and absolutely nothing to worry about. *Zzzzzzzz*.

What is a 'virus' REALLY?

Dr Tom Cowan is one of many contesting the very existence of viruses let alone that they cause disease. This is understandable when there is no scientific evidence for a disease-causing 'virus'. German virologist Dr Stefan Lanka won a landmark case in 2017 in the German Supreme Court over his contention that there is no such thing as a measles virus. He had offered a big prize for anyone who could prove there is and Lanka won his case when someone sought to claim the money. There is currently a prize of more than 225,000 euros on offer from an Isolate Truth Fund for anyone who can prove the isolation of SARS-CoV-2 and its genetic substance. Lanka wrote in an article headed 'The Misconception Called Virus' that scientists think a 'virus' is causing tissue to become diseased and degraded when in fact it is the *processes they are using* which do that – not a 'virus'. Lanka has done an important job in making this point clear as Cowan did in his analysis of the CDC paper. Lanka says that all claims about viruses as disease-causing pathogens are wrong and based on 'easily recognisable, understandable and verifiable misinterpretations.' Scientists believed they were working with 'viruses' in their laboratories when they were really working with 'typical particles of specific dying tissues or cells …' Lanka said that the tissue decaying process claimed to be caused by a 'virus' still happens when no alleged 'virus' is involved. It's the *process* that does the damage and not a 'virus'. The genetic sample is deprived of nutrients, removed from its energy supply through removal from the body and then doused in toxic antibiotics to remove any bacteria. He confirms again that establishment scientists do not (pinch me)

conduct control experiments to see if this is the case and if they did they would see the claims that 'viruses' are doing the damage is nonsense. He adds that during the measles 'virus' court case he commissioned an independent laboratory to perform just such a control experiment and the result was that the tissues and cells died in the exact same way as with alleged 'infected' material. This is supported by a gathering number of scientists, doctors and researchers who reject what is called 'germ theory' or the belief in the body being infected by contagious sources emitted by other people. Researchers Dawn Lester and David Parker take the same stance in their highly-detailed and sourced book *What Really Makes You Ill – Why everything you thought you knew about disease is wrong* which was recommended to me by a number of medical professionals genuinely seeking the truth. Lester and Parker say there is no provable scientific evidence to show that a 'virus' can be transmitted between people or people and animals or animals and people:

> The definition also claims that viruses are the cause of many diseases, as if this has been definitively proven. But this is not the case; there is no original scientific evidence that definitively demonstrates that any virus is the cause of any disease. The burden of proof for any theory lies with those who proposed it; but none of the existing documents provides 'proof' that supports the claim that 'viruses' are pathogens.

Dr Tom Cowan employs one of his clever analogies to describe the process by which a 'virus' is named as the culprit for a disease when what is called a 'virus' is only material released by cells detoxing themselves from infiltration by chemical or radiation poisoning. The tidal wave of technologically-generated radiation in the 'smart' modern world plus all the toxic food and drink are causing this to happen more than ever. Deluded 'scientists' misread this as a gathering impact of what they wrongly label 'viruses'.

Paper can infect houses

Cowan said in an article for davidicke.com – with his tongue only mildly in his cheek – that he believed he had made a tremendous

discovery that may revolutionise science. He had discovered that small bits of paper are alive, 'well alive-ish', can 'infect' houses, and then reproduce themselves inside the house. The result was that this explosion of growth in the paper inside the house causes the house to explode, blowing it to smithereens. His evidence for this new theory is that in the past months he had carefully examined many of the houses in his neighbourhood and found almost no scraps of paper on the lawns and surrounds of the house. There was an occasional stray label, but nothing more. Then he would return to these same houses a week or so later and with a few, not all of them, particularly the old and decrepit ones, he found to his shock and surprise they were littered with stray bits of paper. He knew then that the paper had infected these houses, made copies of itself, and blew up the house. A young boy on a bicycle at one of the sites told him he had seen a demolition crew using dynamite to explode the house the previous week, but Cowan dismissed this as the idle thoughts of silly boys because 'I was on to something big'. He was on to how 'scientists' mistake genetic material in the detoxifying process for something they call a 'virus'. Cowan said of his house and paper story:

> If this sounds crazy to you, it's because it should. This scenario is obviously nuts. But consider this admittedly embellished, for effect, current viral theory that all scientists, medical doctors and virologists currently believe.

He takes the example of the 'novel SARS-Cov2' virus to prove the point. First they take someone with an undefined illness called 'Covid-19' and don't even attempt to find any virus in their sputum. Never mind the scientists still describe how this 'virus', which they have not located attaches to a cell receptor, injects its genetic material, in 'Covid's' case, RNA, into the cell. The RNA once inserted exploits the cell to reproduce itself and makes 'thousands, nay millions, of copies of itself … Then it emerges victorious to claim its next victim':

If you were to look in the scientific literature for proof, actual scientific proof, that uniform SARS-CoV2 viruses have been properly isolated from the sputum of a sick person, that actual spike proteins could be seen protruding from the virus (which has not been found), you would find that such evidence doesn't exist.

If you go looking in the published scientific literature for actual pictures, proof, that these spike proteins or any viral proteins are ever attached to any receptor embedded in any cell membrane, you would also find that no such evidence exists. If you were to look for a video or documented evidence of the intact virus injecting its genetic material into the body of the cell, reproducing itself and then emerging victorious by budding off the cell membrane, you would find that no such evidence exists.

The closest thing you would find is electron micrograph pictures of cellular particles, possibly attached to cell debris, both of which to be seen were stained by heavy metals, a process that completely distorts their architecture within the living organism. This is like finding bits of paper stuck to the blown-up bricks, thereby proving the paper emerged by taking pieces of the bricks on its way out.

The Enders baloney

Cowan describes the 'Covid' story as being just as make-believe as his paper story and he charts back this fantasy to a Nobel Prize winner called John Enders (1897-1985), an American biomedical scientist who has been dubbed 'The Father of Modern Vaccines'. Enders is claimed to have 'discovered' the process of the viral culture which 'proved' that a 'virus' caused measles. Cowan explains how Enders did this 'by using the EXACT same procedure that has been followed by every virologist to find and characterize every new virus since 1954'. Enders took throat swabs from children with measles and immersed them in 2ml of milk. Penicillin (100u/ml) and the antibiotic streptomycin (50,g/ml) were added and the whole mix was centrifuged – rotated at high speed to separate large cellular debris from small particles and molecules as with milk and cream, for example. Cowan says that if the aim is to find little particles of genetic material ('viruses') in the snot from children with measles it would seem that the last thing you would do is mix the snot with other material – milk –that also has genetic material. 'How are you ever going to know whether whatever you found came from the snot or the milk?' He points out that streptomycin is a 'nephrotoxic' or poisonous-to-the-kidney drug. You will see the relevance of that

shortly. Cowan says that it gets worse, much worse, when Enders describes the culture medium upon which the virus 'grows': 'The culture medium consisted of bovine amniotic fluid (90%), beef embryo extract (5%), horse serum (5%), antibiotics and phenol red as an indicator of cell metabolism.' Cowan asks incredulously: 'Did he just say that the culture medium also contained fluids and tissues that are themselves rich sources of genetic material?' The genetic cocktail, or 'medium', is inoculated onto tissue and cells from rhesus monkey *kidney* tissue. This is where the importance of streptomycin comes in and currently-used antimicrobials and other drugs that are *poisonous to kidneys* and used in ALL modern viral cultures (e.g. gentamicin, streptomycin, and amphotericin). Cowan asks: 'How are you ever going to know from this witch's brew where any genetic material comes from as we now have five different sources of rich genetic material in our mix?' Remember, he says, that all genetic material, whether from monkey kidney tissues, bovine serum, milk, etc., is made from the exact same components. The same central question returns: 'How are you possibly going to know that it was the virus that killed the kidney tissue and not the toxic antibiotic and starvation rations on which you are growing the tissue?' John Enders answered the question himself – *you can't*:

A second agent was obtained from an uninoculated culture of monkey kidney cells. The cytopathic changes [death of the cells] it induced in the unstained preparations could not be distinguished with confidence from the viruses isolated from measles.

The death of the cells ('cytopathic changes') happened in exactly the same manner, whether they inoculated the kidney tissue with the measles snot or not, Cowan says. 'This is evidence that the destruction of the tissue, the very proof of viral causation of illness, was not caused by anything in the snot because they saw the same destructive effect when the snot was not even used … the cytopathic, i.e., cell-killing, changes come from the process of the culture itself, not from any virus in any snot, period.' Enders quotes in his 1957 paper a virologist called Ruckle as reporting similar findings 'and in addition has isolated an agent from monkey kidney tissue that is so

far indistinguishable from human measles virus'. In other words, Cowan says, these particles called 'measles viruses' are simply and clearly breakdown products of the starved and poisoned tissue. For measles 'virus' see all 'viruses' including the so-called 'Covid virus'. Enders, the 'Father of Modern Vaccines', also said:

There is a potential risk in employing cultures of primate cells for the production of vaccines composed of attenuated virus, since the presence of other agents possibly latent in primate tissues cannot be definitely excluded by any known method.

Cowan further quotes from a paper published in the journal *Viruses* in May, 2020, while the 'Covid pandemic' was well underway in the media if not in reality. 'EVs' here refers to particles of genetic debris from our own tissues, such as exosomes of which more in a moment: 'The remarkable resemblance between EVs and viruses has caused quite a few problems in the studies focused on the analysis of EVs released during viral infections.' Later the paper adds that to date a reliable method that can actually guarantee a complete separation (of EVs from viruses) DOES NOT EXIST. This was published at a time when a fairy tale 'virus' was claimed in total certainty to be causing a fairy tale 'viral disease' called 'Covid-19' – a fairy tale that was already well on the way to transforming human society in the image that the Cult has worked to achieve for so long. Cowan concludes his article:

To summarize, there is no scientific evidence that pathogenic viruses exist. What we think of as 'viruses' are simply the normal breakdown products of dead and dying tissues and cells. When we are well, we make fewer of these particles; when we are starved, poisoned, suffocated by wearing masks, or afraid, we make more.

There is no engineered virus circulating and making people sick. People in laboratories all over the world are making genetically modified products to make people sick. These are called vaccines. There is no virome, no 'ecosystem' of viruses, viruses are not 8%, 50% or 100 % of our genetic material. These are all simply erroneous ideas based on the misconception called a virus.

What is 'Covid'? Load of bollocks

The background described here by Cowan and Lanka was emphasised in the first video presentation that I saw by Dr Andrew Kaufman when he asked whether the 'Covid virus' was in truth a natural defence mechanism of the body called 'exosomes'. These are released by cells when in states of toxicity – see the same themes returning over and over. They are released ever more profusely as chemical and radiation toxicity increases and think of the potential effect therefore of 5G alone as its destructive frequencies infest the human energetic information field with a gathering pace (5G went online in Wuhan in 2019 as the 'virus' emerged). I'll have more about this later. Exosomes transmit a warning to the rest of the body that 'Houston, we have a problem'. Kaufman presented images of exosomes and compared them with 'Covid' under an electron microscope and the similarity was remarkable. They both attach to the same cell receptors (*claimed* in the case of 'Covid'), contain the same genetic material in the form of RNA or ribonucleic acid, and both are found in 'viral cell cultures' with damaged or dying cells. James Hildreth MD, President and Chief Executive Officer of the Meharry Medical College at Johns Hopkins, said: 'The virus is fully an exosome in every sense of the word.' Kaufman's conclusion was that there is no 'virus': 'This entire pandemic is a completely manufactured crisis … there is no evidence of anyone dying from [this] illness.' Dr Tom Cowan and Sally Fallon Morell, authors of *The Contagion Myth*, published a statement with Dr Kaufman in February, 2021, explaining why the 'virus' does not exist and you can read it that in full in the Appendix.

'Virus' theory can be traced to the 'cell theory' in 1858 of German physician Rudolf Virchow (1821-1920) who contended that disease originates from a single cell infiltrated by a 'virus'. Dr Stefan Lanka said that findings and insights with respect to the structure, function and central importance of tissues in the creation of life, which were already known in 1858, comprehensively refute the cell theory. Virchow ignored them. We have seen the part later played by John Enders in the 1950s and Lanka notes that infection theories were only established as a global dogma through the policies and

eugenics of the Third Reich in Nazi Germany (creation of the same Sabbatian cult behind the 'Covid' hoax). Lanka said: 'Before 1933, scientists dared to contradict this theory; after 1933, these critical scientists were silenced'. Dr Tom Cowan's view is that ill-heath is caused by too much of something, too little of something, or toxification from chemicals and radiation – not contagion. We must also highlight as a major source of the 'virus' theology a man still called the 'Father of Modern Virology' – Thomas Milton Rivers (1888-1962). There is no way given the Cult's long game policy that it was a coincidence for the 'Father of Modern Virology' to be director of the Rockefeller Institute for Medical Research from 1937 to 1956 when he is credited with making the Rockefeller Institute a leader in 'viral research'. Cult Rockefellers were the force behind the creation of Big Pharma 'medicine', established the World Health Organisation in 1948, and have long and close associations with the Gates family that now runs the WHO during the pandemic hoax through mega-rich Cult gofer and psychopath Bill Gates.

Only a Renegade Mind can see through all this bullshit by asking the questions that need to be answered, not taking 'no' or prevarication for an answer, and certainly not hiding from the truth in fear of speaking it. Renegade Minds have always changed the world for the better and they will change this one no matter how bleak it may currently appear to be.

CHAPTER SIX

Sequence of deceit

If you tell the truth, you don't have to remember anything
Mark Twain

Against the background that I have laid out this far the sequence that took us from an invented 'virus' in Cult-owned China in late 2019 to the fascist transformation of human society can be seen and understood in a whole new context.

We were told that a deadly disease had broken out in Wuhan and the world media began its campaign (coordinated by behavioural psychologists as we shall see) to terrify the population into unquestioning compliance. We were shown images of Chinese people collapsing in the street which never happened in the West with what was supposed to be the same condition. In the earliest days when alleged cases and deaths were few the fear register was hysterical in many areas of the media and this would expand into the common media narrative across the world. The real story was rather different, but we were never told that. The Chinese government, one of the Cult's biggest centres of global operation, said they had discovered a new illness with flu-like and pneumonia-type symptoms in a city with such toxic air that it is overwhelmed with flu-like symptoms, pneumonia and respiratory disease. Chinese scientists said it was a new – 'novel' – coronavirus which they called Sars-Cov-2 and that it caused a disease they labelled 'Covid-19'. There was no evidence for this and the 'virus' has never to this day been isolated, purified and its genetic code established from that. It

was from the beginning a computer-generated fiction. Stories of Chinese whistleblowers saying the number of deaths was being supressed or that the 'new disease' was related to the Wuhan bio-lab misdirected mainstream and alternative media into cul-de-sacs to obscure the real truth – there was no 'virus'.

Chinese scientists took genetic material from the lung fluid of just a few people and said they had found a 'new' disease when this material had a wide range of content. There was no evidence for a 'virus' for the very reasons explained in the last two chapters. The 'virus' has never been shown to (a) exist and (b) cause any disease. People were diagnosed on symptoms that are so widespread in Wuhan and polluted China and with a PCR test that can't detect infectious disease. On this farce the whole global scam was sold to the rest of the world which would also diagnose respiratory disease as 'Covid-19' from symptoms alone or with a PCR test not testing for a 'virus'. Flu miraculously disappeared *worldwide* in 2020 and into 2021 as it was redesignated 'Covid-19'. It was really the same old flu with its 'flu-like' symptoms attributed to 'flu-like' 'Covid-19'. At the same time with very few exceptions the Chinese response of draconian lockdown and fascism was the chosen weapon to respond across the West as recommended by the Cult-owned Tedros at the Cult-owned World Health Organization run by the Cult-owned Gates. All was going according to plan. Chinese scientists – everything in China is controlled by the Cult-owned government – compared their contaminated RNA lung-fluid material with other RNA sequences and said it appeared to be just under 80 percent identical to the SARS-CoV-1 'virus' claimed to be the cause of the SARS (severe acute respiratory syndrome) 'outbreak' in 2003. They decreed that because of this the 'new virus' had to be related and they called it SARS-CoV-2. There are some serious problems with this assumption and *assumption* was all it was. Most 'factual' science turns out to be assumptions repeated into everyone-knows-that. A match of under 80-percent is meaningless. Dr Kaufman makes the point that there's a *96 percent* genetic correlation between humans and chimpanzees, but 'no one would say our genetic material is part

of the chimpanzee family'. Yet the Chinese authorities were claiming that a much lower percentage, less than 80 percent, proved the existence of a new 'coronavirus'. For goodness sake human DNA is 60 percent similar to a *banana*.

You are feeling sleepy

The entire 'Covid' hoax is a global Psyop, a psychological operation to program the human mind into believing and fearing a complete fantasy. A crucial aspect of this was what *appeared* to happen in Italy. It was all very well streaming out daily images of an alleged catastrophe in Wuhan, but to the Western mind it was still on the other side of the world in a very different culture and setting. A reaction of 'this could happen to me and my family' was still nothing like as intense enough for the mind-doctors. The Cult needed a Western example to push people over that edge and it chose Italy, one of its major global locations going back to the Roman Empire. An Italian 'Covid' crisis was manufactured in a particular area called Lombardy which just happens to be notorious for its toxic air and therefore respiratory disease. Wuhan, China, déjà vu. An hysterical media told horror stories of Italians dying from 'Covid' in their droves and how Lombardy hospitals were being overrun by a tidal wave of desperately ill people needing treatment after being struck down by the 'deadly virus'. Here was the psychological turning point the Cult had planned. Wow, if this is happening in Italy, the Western mind concluded, this indeed could happen to me and my family. Another point is that Italian authorities responded by following the Chinese blueprint so vehemently recommended by the Cult-owned World Health Organization. They imposed fascistic lockdowns on the whole country viciously policed with the help of surveillance drones sweeping through the streets seeking out anyone who escaped from mass house arrest. Livelihoods were destroyed and psychology unravelled in the way we have witnessed since in all lockdown countries. Crucial to the plan was that Italy responded in this way to set the precedent of suspending freedom and imposing fascism in a 'Western liberal democracy'. I emphasised in an

animated video explanation on davidicke.com posted in the summer of 2020 how important it was to the Cult to expand the Chinese lockdown model across the West. Without this, and the bare-faced lie that non-symptomatic people could still transmit a 'disease' they didn't have, there was no way locking down the whole population, sick and not sick, could be pulled off. At just the right time and with no evidence Cult operatives and gofers claimed that people without symptoms could pass on the 'disease'. In the name of protecting the 'vulnerable' like elderly people, who lockdowns would kill by the tens of thousands, we had for the first time healthy people told to isolate as well as the sick. The great majority of people who tested positive had no symptoms because there was nothing wrong with them. It was just a trick made possible by a test not testing for the 'virus'.

Months after my animated video the Gates-funded Professor Neil Ferguson at the Gates-funded Imperial College confirmed that I was right. He didn't say it in those terms, naturally, but he did say it. Ferguson will enter the story shortly for his outrageously crazy 'computer models' that led to Britain, the United States and many other countries following the Chinese and now Italian methods of response. Put another way, following the Cult script. Ferguson said that SAGE, the UK government's scientific advisory group which has controlled 'Covid' policy from the start, wanted to follow the Chinese lockdown model (while they all continued to work and be paid), but they wondered if they could possibly, in Ferguson's words, 'get away with it in Europe'. 'Get away with it'? Who the hell do these moronic, arrogant people think they are? This appalling man Ferguson said that once Italy went into national lockdown they realised they, too, could mimic China:

> It's a communist one-party state, we said. We couldn't get away with it in Europe, we thought … and then Italy did it. And we realised we could. Behind this garbage from Ferguson is a simple fact: Doing the same as China in every country was the plan from the start and Ferguson's 'models' would play a central role in achieving that. It's just a coincidence, of course, and absolutely nothing to worry your little head about.

Oops, sorry, our mistake

Once the Italian segment of the Psyop had done the job it was designed to do a very different story emerged. Italian authorities revealed that 99 percent of those who had 'died from Covid-19' in Italy had one, two, three, or more 'co-morbidities' or illnesses and health problems that could have ended their life. The US Centers for Disease Control and Prevention (CDC) published a figure of 94 percent for Americans dying of 'Covid' while having other serious medical conditions – on average two to three (some five or six) other potential causes of death. In terms of death from an unproven 'virus' I say it is 100 percent. The other one percent in Italy and six percent in the US would presumably have died from 'Covid's' flu-like symptoms with a range of other possible causes in conjunction with a test not testing for the 'virus'. Fox News reported that even more startling figures had emerged in one US county in which 410 of 422 deaths attributed to 'Covid-19' had other potentially deadly health conditions. The Italian National Health Institute said later that the average age of people dying with a 'Covid-19' diagnosis in Italy was about 81. Ninety percent were over 70 with ten percent over 90. In terms of other reasons to die some 80 percent had two or more chronic diseases with half having three or more including cardiovascular problems, diabetes, respiratory problems and cancer. Why is the phantom 'Covid-19' said to kill overwhelmingly old people and hardly affect the young? Old people continually die of many causes and especially respiratory disease which you can re-diagnose 'Covid-19' while young people die in tiny numbers by comparison and rarely of respiratory disease. Old people 'die of Covid' because they die of other things that can be redesignated 'Covid' and it really is that simple.

Flu has flown

The blueprint was in place. Get your illusory 'cases' from a test not testing for the 'virus' and redesignate other causes of death as 'Covid-19'. You have an instant 'pandemic' from something that is nothing more than a computer-generated fiction. With near-on a

billion people having 'flu-like' symptoms every year the potential was limitless and we can see why flu quickly and apparently miraculously disappeared *worldwide* by being diagnosed 'Covid-19'. The painfully bloody obvious was explained away by the childlike media in headlines like this in the UK '*Independent*': 'Not a single case of flu detected by Public Health England this year as Covid restrictions suppress virus'. I kid you not. The masking, social distancing and house arrest that did not make the 'Covid virus' disappear somehow did so with the 'flu virus'. Even worse the article, by a bloke called Samuel Lovett, suggested that maybe the masking, sanitising and other 'Covid' measures should continue to keep the flu away. With a ridiculousness that disturbs your breathing (it's 'Covid-19') the said Lovett wrote: 'With widespread social distancing and mask-wearing measures in place throughout the UK, the usual routes of transmission for influenza have been blocked.' He had absolutely no evidence to support that statement, but look at the consequences of him acknowledging the obvious. With flu not disappearing at all and only being relabelled 'Covid-19' he would have to contemplate that 'Covid' was a hoax on a scale that is hard to imagine. You need guts and commitment to truth to even go there and that's clearly something Samuel Lovett does not have in abundance. He would never have got it through the editors anyway.

Tens of thousands die in the United States alone every winter from flu including many with pneumonia complications. CDC figures record *45 million* Americans diagnosed with flu in 2017-2018 of which 61,000 died and some reports claim 80,000. Where was the same hysteria then that we have seen with 'Covid-19'? Some 250,000 Americans are admitted to hospital with pneumonia every year with about 50,000 cases proving fatal. About 65 million suffer respiratory disease every year and three million deaths makes this the third biggest cause of death worldwide. You only have to redesignate a portion of all these people 'Covid-19' and you have an instant global pandemic or the *appearance* of one. Why would doctors do this? They are told to do this and all but a few dare not refuse those who must be obeyed. Doctors in general are not researching their own

knowledge and instead take it direct and unquestioned from the authorities that own them and their careers. The authorities say they must now diagnose these symptoms 'Covid-19' and not flu, or whatever, and they do it. Dark suits say put 'Covid-19' on death certificates no matter what the cause of death and the doctors do it. Renegade Minds don't fall for the illusion that doctors and medical staff are all highly-intelligent, highly-principled, seekers of medical truth. *Some are*, but not the majority. They are repeaters, gofers, and yes sir, no sir, purveyors of what the system demands they purvey. The 'Covid' con is not merely confined to diseases of the lungs. Instructions to doctors to put 'Covid-19' on death certificates for anyone dying of *anything* within 28 days (or much more) of a positive test not testing for the 'virus' opened the floodgates. The term dying *with* 'Covid' and not *of* 'Covid' was coined to cover the truth. Whether it was a *with* or an *of* they were all added to the death numbers attributed to the 'deadly virus' compiled by national governments and globally by the Gates-funded Johns Hopkins operation in the United States that was so involved in those 'pandemic' simulations. Fraudulent deaths were added to the ever-growing list of fraudulent 'cases' from false positives from a false test. No wonder Professor Walter Ricciardi, scientific advisor to the Italian minister of health, said after the Lombardy hysteria had done its job that 'Covid' death rates were due to Italy having the second oldest population in the world and to *how hospitals record deaths*:

The way in which we code deaths in our country is very generous in the sense that all the people who die in hospitals with the coronavirus are deemed to be dying of the coronavirus. On re-evaluation by the National Institute of Health, only 12 per cent of death certificates have shown a direct causality from coronavirus, while 88 per cent of patients who have died have at least one pre-morbidity – many had two or three.

This is extraordinary enough when you consider the propaganda campaign to use Italy to terrify the world, but how can they even say twelve percent were genuine when the 'virus' has not been shown to exist, its 'code' is a computer program, and diagnosis comes from a test not testing for it? As in China, and soon the world, 'Covid-19' in

Italy was a redesignation of diagnosis. Lies and corruption were to become the real 'pandemic' fuelled by a pathetically-compliant medical system taking its orders from the tiny few at the top of their national hierarchy who answered to the World Health Organization which answers to Gates and the Cult. Doctors were told – ordered – to diagnose a particular set of symptoms 'Covid-19' and put that on the death certificate for any cause of death if the patient had tested positive with a test not testing for the virus or had 'Covid' symptoms like the flu. The United States even introduced big financial incentives to manipulate the figures with hospitals receiving £4,600 from the Medicare system for diagnosing someone with regular pneumonia, $13,000 if they made the diagnosis from the same symptoms 'Covid-19' pneumonia, and $39, 000 if they put a 'Covid' diagnosed patient on a ventilator that would almost certainly kill them. A few – painfully and pathetically few – medical whistleblowers revealed (before Cult-owned YouTube deleted their videos) that they had been instructed to 'let the patient crash' and put them straight on a ventilator instead of going through a series of far less intrusive and dangerous methods as they would have done before the pandemic hoax began and the financial incentives kicked in. We are talking cold-blooded murder given that ventilators are so damaging to respiratory systems they are usually the last step before heaven awaits. Renegade Minds never fall for the belief that people in white coats are all angels of mercy and cannot be full-on psychopaths. I have explained in detail in *The Answer* how what I am describing here played out across the world coordinated by the World Health Organization through the medical hierarchies in almost every country.

Medical scientist calls it

Information about the non-existence of the 'virus' began to emerge for me in late March, 2020, and mushroomed after that. I was sent an email by Sir Julian Rose, a writer, researcher, and organic farming promotor, from a medical scientist friend of his in the United States. Even at that early stage in March the scientist was able to explain

how the 'Covid' hoax was being manipulated. He said there were no reliable tests for a specific 'Covid-19 virus' and nor were there any reliable agencies or media outlets for reporting numbers of actual 'Covid-19' cases. We have seen in the long period since then that he was absolutely right. 'Every action and reaction to Covid-19 is based on totally flawed data and we simply cannot make accurate assessments,' he said. Most people diagnosed with 'Covid-19' were showing nothing more than cold and flu-like symptoms 'because most coronavirus strains *are* nothing more than cold/flu-like symptoms'. We had farcical situations like an 84-year-old German man testing positive for 'Covid-19' and his nursing home ordered to quarantine only for him to be found to have a common cold. The scientist described back then why PCR tests and what he called the 'Mickey Mouse test kits' were useless for what they were claimed to be identifying. 'The idea these kits can isolate a specific virus like Covid-19 is nonsense,' he said. Significantly, he pointed out that 'if you want to create a totally false panic about a totally false pandemic – pick a coronavirus'. This is exactly what the Cult-owned Gates, World Economic Forum and Johns Hopkins University did with their Event 201 'simulation' followed by their real-life simulation called the 'pandemic'. The scientist said that all you had to do was select the sickest of people with respiratory-type diseases in a single location – 'say Wuhan' – and administer PCR tests to them. You can then claim that anyone showing 'viral sequences' similar to a coronavirus 'which will inevitably be quite a few' is suffering from a 'new' disease:

> Since you already selected the sickest flu cases a fairly high proportion of your sample will go on to die. You can then say this 'new' virus has a CFR [case fatality rate] higher than the flu and use this to infuse more concern and do more tests which will of course produce more 'cases', which expands the testing, which produces yet more 'cases' and so on and so on. Before long you have your 'pandemic', and all you have done is use a simple test kit trick to convert the worst flu and pneumonia cases into something new that doesn't ACTUALLY EXIST [my emphasis].

He said that you then 'just run the same scam in other countries' and make sure to keep the fear message running high 'so that people

will feel panicky and less able to think critically'. The only problem to overcome was the fact *there is no* actual new deadly pathogen and only regular sick people. This meant that deaths from the 'new deadly pathogen' were going to be way too low for a real new deadly virus pandemic, but he said this could be overcome in the following ways – all of which would go on to happen:

1. You can claim this is just the beginning and more deaths are imminent [you underpin this with fantasy 'computer projections']. Use this as an excuse to quarantine everyone and then claim the quarantine prevented the expected millions of dead.

2. You can [say that people] 'minimizing' the dangers are irresponsible and bully them into not talking about numbers.

3. You can talk crap about made up numbers hoping to blind people with pseudoscience.

4. You can start testing well people (who, of course, will also likely have shreds of coronavirus [RNA] in them) and thus inflate your 'case figures' with 'asymptomatic carriers' (you will of course have to spin that to sound deadly even though any virologist knows the more symptom-less cases you have the less deadly is your pathogen).

The scientist said that if you take these simple steps 'you can have your own entirely manufactured pandemic up and running in weeks'. His analysis made so early in the hoax was brilliantly prophetic of what would actually unfold. Pulling all the information together in these recent chapters we have this is simple 1, 2, 3, of how you can delude virtually the entire human population into believing in a 'virus' that doesn't exist:

- A 'Covid case' is someone who tests positive with a test not testing for the 'virus'.

- A 'Covid death' is someone who dies of *any cause* within 28 days (or much longer) of testing positive with a test not testing for the 'virus.

- Asymptomatic means there is nothing wrong with you, but they claim you can pass on what you don't have to justify locking

down (quarantining) healthy people in totality.

The foundations of the hoax are that simple. A study involving ten million people in Wuhan, published in November, 2020, demolished the whole lie about those without symptoms passing on the 'virus'. They found '300 asymptomatic cases' and traced their contacts to find that not one of them was detected with the 'virus'. 'Asymptomatic' patients and their contacts were isolated for no less than two weeks and nothing changed. I know it's all crap, but if you are going to claim that those without symptoms can transmit 'the virus' then you must produce evidence for that and they never have. Even World Health Organization official Dr Maria Van Kerkhove, head of the emerging diseases and zoonosis unit, said as early as June, 2020, that she doubted the validity of asymptomatic transmission. She said that 'from the data we have, it still seems to be rare that an asymptomatic person actually transmits onward to a secondary individual' and by 'rare' she meant that she couldn't cite any case of asymptomatic transmission.

The Ferguson factor

The problem for the Cult as it headed into March, 2020, when the script had lockdown due to start, was that despite all the manipulation of the case and death figures they still did not have enough people alleged to have died from 'Covid' to justify mass house arrest. This was overcome in the way the scientist described: 'You can claim this is just the beginning and more deaths are imminent … Use this as an excuse to quarantine everyone and then claim the quarantine prevented the expected millions of dead.' Enter one Professor Neil Ferguson, the Gates-funded 'epidemiologist' at the Gates-funded Imperial College in London. Ferguson is Britain's Christian Drosten in that he has a dire record of predicting health outcomes, but is still called upon to advise government on the next health outcome when another 'crisis' comes along. This may seem to be a strange and ridiculous thing to do. Why would you keep turning for policy guidance to people who have a history of being

monumentally wrong? Ah, but it makes sense from the Cult point of view. These 'experts' keep on producing predictions that suit the Cult agenda for societal transformation and so it was with Neil Ferguson as he revealed his horrific (and clearly insane) computer model predictions that allowed lockdowns to be imposed in Britain, the United States and many other countries. Ferguson does not have even an A-level in biology and would appear to have no formal training in computer modelling, medicine or epidemiology, according to Derek Winton, an MSc in Computational Intelligence. He wrote an article somewhat aghast at what Ferguson did which included taking no account of respiratory disease 'seasonality' which means it is far worse in the winter months. Who would have thought that respiratory disease could be worse in the winter? Well, certainly not Ferguson.

The massively China-connected Imperial College and its bizarre professor provided the excuse for the long-incubated Chinese model of human control to travel westward at lightning speed. Imperial College confirms on its website that it collaborates with the Chinese Research Institute; publishes more than 600 research papers every year with Chinese research institutions; has 225 Chinese staff; 2,600 Chinese students – the biggest international group; 7,000 former students living in China which is the largest group outside the UK; and was selected for a tour by China's President Xi Jinping during his state visit to the UK in 2015. The college takes major donations from China and describes itself as the UK's number one university collaborator with Chinese research institutions. The China communist/fascist government did not appear phased by the woeful predictions of Ferguson and Imperial when during the lockdown that Ferguson induced the college signed a five-year collaboration deal with China tech giant Huawei that will have Huawei's indoor 5G network equipment installed at the college's West London tech campus along with an 'AI cloud platform'. The deal includes Chinese sponsorship of Imperial's Venture Catalyst entrepreneurship competition. Imperial is an example of the enormous influence the Chinese government has within British and North American

universities and research centres – and further afield. Up to 200 academics from more than a dozen UK universities are being investigated on suspicion of 'unintentionally' helping the Chinese government build weapons of mass destruction by 'transferring world-leading research in advanced military technology such as aircraft, missile designs and cyberweapons'. Similar scandals have broken in the United States, but it's all a coincidence. Imperial College serves the agenda in many other ways including the promotion of every aspect of the United Nations Agenda 21/2030 (the Great Reset) and produced computer models to show that human-caused 'climate change' is happening when in the real world it isn't. Imperial College is driving the climate agenda as it drives the 'Covid' agenda (both Cult hoaxes) while Patrick Vallance, the UK government's Chief Scientific Adviser on 'Covid', was named Chief Scientific Adviser to the UN 'climate change' conference known as COP26 hosted by the government in Glasgow, Scotland. 'Covid' and 'climate' are fundamentally connected.

Professor Woeful

From Imperial's bosom came Neil Ferguson still advising government despite his previous disasters and it was announced early on that he and other key people like UK Chief Medical Adviser Chris Whitty had caught the 'virus' as the propaganda story was being sold. Somehow they managed to survive and we had Prime Minister Boris Johnson admitted to hospital with what was said to be a severe version of the 'virus' in this same period. His whole policy and demeanour changed when he returned to Downing Street. It's a small world with these government advisors – especially in their communal connections to Gates – and Ferguson had partnered with Whitty to write a paper called 'Infectious disease: Tough choices to reduce Ebola transmission' which involved another scare-story that didn't happen. Ferguson's 'models' predicted that up to150, 000 could die from 'mad cow disease', or BSE, and its version in sheep if it was transmitted to humans. BSE was not transmitted and instead triggered by an organophosphate pesticide used to treat a pest on

cows. Fewer than 200 deaths followed from the human form. Models by Ferguson and his fellow incompetents led to the unnecessary culling of millions of pigs, cattle and sheep in the foot and mouth outbreak in 2001 which destroyed the lives and livelihoods of farmers and their families who had often spent decades building their herds and flocks. Vast numbers of these animals did not have foot and mouth and had no contact with the infection. Another 'expert' behind the cull was Professor Roy Anderson, a computer modeller at Imperial College specialising in the epidemiology of *human*, not animal, disease. Anderson has served on the Bill and Melinda Gates Grand Challenges in Global Health advisory board and chairs another Gates-funded organisation. Gates is everywhere.

In a precursor to the 'Covid' script Ferguson backed closing schools 'for prolonged periods' over the swine flu 'pandemic' in 2009 and said it would affect a third of the world population if it continued to spread at the speed he claimed to be happening. His mates at Imperial College said much the same and a news report said: 'One of the authors, the epidemiologist and disease modeller Neil Ferguson, who sits on the World Health Organisation's emergency committee for the outbreak, said the virus had "full pandemic potential".' Professor Liam Donaldson, the Chris Whitty of his day as Chief Medical Officer, said the worst case could see 30 percent of the British people infected by swine flu with 65,000 dying. Ferguson and Donaldson were indeed proved correct when at the end of the year the number of deaths attributed to swine flu was 392. The term 'expert' is rather liberally applied unfortunately, not least to complete idiots. Swine flu 'projections' were great for GlaxoSmithKline (GSK) as millions rolled in for its Pandemrix influenza vaccine which led to brain damage with children most affected. The British government (taxpayers) paid out more than £60 million in compensation after GSK was given immunity from prosecution. Yet another 'Covid' déjà vu. Swine flu was supposed to have broken out in Mexico, but Dr Wolfgang Wodarg, a German doctor, former member of parliament and critic of the 'Covid' hoax, observed 'the spread of swine flu' in Mexico City at the time. He

said: 'What we experienced in Mexico City was a very mild flu which did not kill more than usual – which killed even fewer people than usual.' Hyping the fear against all the facts is not unique to 'Covid' and has happened many times before. Ferguson is reported to have over-estimated the projected death toll of bird flu (H5N1) by some three million-fold, but bird flu vaccine makers again made a killing from the scare. This is some of the background to the Neil Ferguson who produced the perfectly-timed computer models in early 2020 predicting that half a million people would die in Britain without draconian lockdown and 2.2 million in the United States. Politicians panicked, people panicked, and lockdowns of alleged short duration were instigated to 'flatten the curve' of cases gleaned from a test not testing for the 'virus'. I said at the time that the public could forget the 'short duration' bit. This was an agenda to destroy the livelihoods of the population and force them into mass control through dependency and there was going to be nothing 'short' about it. American researcher Daniel Horowitz described the consequences of the 'models' spewed out by Gates-funded Ferguson and Imperial College:

> What led our government and the governments of many other countries into panic was a single Imperial College of UK study, funded by global warming activists, that predicted 2.2 million deaths if we didn't lock down the country. In addition, the reported 8-9% death rate in Italy scared us into thinking there was some other mutation of this virus that they got, which might have come here.
>
> Together with the fact that we were finally testing and had the ability to actually report new cases, we thought we were headed for a death spiral. But again … we can't flatten a curve if we don't know when the curve started.

How about it *never* started?

Giving them what they want

An investigation by German news outlet *Welt Am Sonntag (World on Sunday)* revealed how in March, 2020, the German government gathered together 'leading scientists from several research institutes and universities' and 'together, they were to produce a [modelling]

paper that would serve as legitimization for further tough political measures'. The Cult agenda was justified by computer modelling not based on evidence or reality; it was specifically constructed to justify the Cult demand for lockdowns all over the world to destroy the independent livelihoods of the global population. All these modellers and everyone responsible for the 'Covid' hoax have a date with a trial like those in Nuremberg after World War Two when Nazis faced the consequences of their war crimes. These corrupt-beyond-belief 'modellers' wrote the paper according to government instructions and it said that that if lockdown measures were lifted then up to one million Germans would die from 'Covid-19' adding that some would die 'agonizingly at home, gasping for breath' unable to be treated by hospitals that couldn't cope. All lies. No matter – it gave the Cult all that it wanted. What did long-time government 'modeller' Neil Ferguson say? If the UK and the United States didn't lockdown half a million would die in Britain and 2.2 million Americans. Anyone see a theme here? 'Modellers' are such a crucial part of the lockdown strategy that we should look into their background and follow the money. Researcher Rosemary Frei produced an excellent article headlined 'The Modelling-paper Mafiosi'. She highlights a guy called John Edmunds, a British epidemiologist, and professor in the Faculty of Epidemiology and Population Health at the London School of Hygiene & Tropical Medicine. He studied at Imperial College. Edmunds is a member of government 'Covid' advisory bodies which have been dictating policy, the New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG) and the Scientific Advisory Group for Emergencies (SAGE).

Ferguson, another member of NERVTAG and SAGE, led the way with the original 'virus' and Edmunds has followed in the 'variant' stage and especially the so-called UK or Kent variant known as the 'Variant of Concern' (VOC) B.1.1.7. He said in a co-written report for the Centre for Mathematical modelling of Infectious Diseases at the London School of Hygiene and Tropical Medicine, with input from the Centre's 'Covid-19' Working Group, that there was 'a realistic

possibility that VOC B.1.1.7 is associated with an increased risk of death compared to non-VOC viruses'. Fear, fear, fear, get the vaccine, fear, fear, fear, get the vaccine. Rosemary Frei reveals that almost all the paper's authors and members of the modelling centre's 'Covid-19' Working Group receive funding from the Bill and Melinda Gates Foundation and/or the associated Gates-funded Wellcome Trust. The paper was published by e-journal *Medr χiv* which only publishes papers not peer-reviewed and the journal was established by an organisation headed by Facebook's Mark Zuckerberg and his missus. What a small world it is. Frei discovered that Edmunds is on the Scientific Advisory Board of the Coalition for Epidemic Preparedness Innovations (CEPI) which was established by the Bill and Melinda Gates Foundation, Klaus Schwab's Davos World Economic Forum and Big Pharma giant Wellcome. CEPI was 'launched in Davos [in 2017] to develop vaccines to stop future epidemics', according to its website. 'Our mission is to accelerate the development of vaccines against emerging infectious diseases and enable equitable access to these vaccines for people during outbreaks.' What kind people they are. Rosemary Frei reveals that Public Health England (PHE) director Susan Hopkins is an author of her organisation's non-peer-reviewed reports on 'new variants'. Hopkins is a professor of infectious diseases at London's Imperial College which is gifted tens of millions of dollars a year by the Bill and Melinda Gates Foundation. Gates-funded modelling disaster Neil Ferguson also co-authors Public Health England reports and he spoke in December, 2020, about the potential danger of the B.1.1.7. 'UK variant' promoted by Gates-funded modeller John Edmunds. When I come to the 'Covid vaccines' the 'new variants' will be shown for what they are – bollocks.

Connections, connections

All these people and modellers are lockdown-obsessed or, put another way, they demand what the Cult demands. Edmunds said in January, 2021, that to ease lockdowns too soon would be a disaster and they had to 'vaccinate much, much, much more widely than the

elderly'. Rosemary Frei highlights that Edmunds is married to Jeanne Pimenta who is described in a LinkedIn profile as director of epidemiology at GlaxoSmithKline (GSK) and she held shares in the company. Patrick Vallance, co-chair of SAGE and the government's Chief Scientific Adviser, is a former executive of GSK and has a deferred bonus of shares in the company worth £600,000. GSK has serious business connections with Bill Gates and is collaborating with mRNA-'vaccine' company CureVac to make 'vaccines' for the new variants that Edmunds is talking about. GSK is planning a 'Covid vaccine' with drug giant Sanofi. Puppet Prime Minister Boris Johnson announced in the spring of 2021 that up to 60 million vaccine doses were to be made at the GSK facility at Barnard Castle in the English North East. Barnard Castle, with a population of just 6,000, was famously visited in breach of lockdown rules in April, 2020, by Johnson aide Dominic Cummings who said that he drove there 'to test his eyesight' before driving back to London. Cummings would be better advised to test his integrity – not that it would take long. The GSK facility had nothing to do with his visit then although I'm sure Patrick Vallance would have been happy to arrange an introduction and some tea and biscuits. Ruthless psychopath Gates has made yet another fortune from vaccines in collaboration with Big Pharma companies and gushes at the phenomenal profits to be made from vaccines – more than a 20-to-1 return as he told one interviewer. Gates also tweeted in December, 2019, with the foreknowledge of what was coming: 'What's next for our foundation? I'm particularly excited about what the next year could mean for one of the best buys in global health: vaccines.'

Modeller John Edmunds is a big promotor of vaccines as all these people appear to be. He's the dean of the London School of Hygiene & Tropical Medicine's Faculty of Epidemiology and Population Health which is primarily funded by the Bill and Melinda Gates Foundation and the Gates-established and funded GAVI vaccine alliance which is the Gates vehicle to vaccinate the world. The organisation Doctors Without Borders has described GAVI as being 'aimed more at supporting drug-industry desires to promote new

products than at finding the most efficient and sustainable means for fighting the diseases of poverty'. But then that's why the psychopath Gates created it. John Edmunds said in a video that the London School of Hygiene & Tropical Medicine is involved in every aspect of vaccine development including large-scale clinical trials. He contends that mathematical modelling can show that vaccines protect individuals and society. That's on the basis of shit in and shit out, I take it. Edmunds serves on the UK Vaccine Network as does Ferguson and the government's foremost 'Covid' adviser, the grim-faced, dark-eyed Chris Whitty. The Vaccine Network says it works 'to support the government to identify and shortlist targeted investment opportunities for the most promising vaccines and vaccine technologies that will help combat infectious diseases with epidemic potential, and to address structural issues related to the UK's broader vaccine infrastructure'. Ferguson is acting Director of the Imperial College Vaccine Impact Modelling Consortium which has funding from the Bill and Melina Gates Foundation and the Gates-created GAVI 'vaccine alliance'. Anyone wonder why these characters see vaccines as the answer to every problem? Ferguson is wildly enthusiastic in his support for GAVI's campaign to vaccine children en masse in poor countries. You would expect someone like Gates who has constantly talked about the need to reduce the population to want to fund vaccines to keep more people alive. I'm sure that's why he does it. The John Edmunds London School of Hygiene & Tropical Medicine (LSHTM) has a Vaccines Manufacturing Innovation Centre which develops, tests and commercialises vaccines. Rosemary Frei writes:

> The vaccines centre also performs affiliated activities like combating 'vaccine hesitancy'. The latter includes the Vaccine Confidence Project. The project's stated purpose is, among other things, 'to provide analysis and guidance for early response and engagement with the public to ensure sustained confidence in vaccines and immunisation'. The Vaccine Confidence Project's director is LSHTM professor Heidi Larson. For more than a decade she's been researching how to combat vaccine hesitancy.

How the bloody hell can blokes like John Edmunds and Neil Ferguson with those connections and financial ties model 'virus' case

and death projections for the government and especially in a way that gives their paymasters like Gates exactly what they want? It's insane, but this is what you find throughout the world.

'Covid' is not dangerous, oops, wait, yes it is

Only days before Ferguson's nightmare scenario made Jackboot Johnson take Britain into a China-style lockdown to save us from a deadly 'virus' the UK government website gov.uk was reporting something very different to Ferguson on a page of official government guidance for 'high consequence infectious diseases (HCID)'. It said this about 'Covid-19':

As of 19 March 2020, COVID-19 *is no longer considered to be a high consequence infectious diseases (HCID) in the UK* [my emphasis]. The 4 nations public health HCID group made an interim recommendation in January 2020 to classify COVID-19 as an HCID. This was based on consideration of the UK HCID criteria about the virus and the disease with information available during the early stages of the outbreak.

Now that more is known about COVID-19, the public health bodies in the UK have reviewed the most up to date information about COVID-19 against the UK HCID criteria. They have determined that several features have now changed; in particular, more information is available about mortality rates (low overall), and there is now greater clinical awareness and a specific and sensitive laboratory test, the availability of which continues to increase. The Advisory Committee on Dangerous Pathogens (ACDP) is also of the opinion that COVID-19 should no longer be classified as an HCID.

Soon after the government had been exposed for downgrading the risk they upgraded it again and everyone was back to singing from the same Cult hymn book. Ferguson and his fellow Gates clones indicated that lockdowns and restrictions would have to continue until a Gates-funded vaccine was developed. Gates said the same because Ferguson and his like were repeating the Gates script which is the Cult script. 'Flatten the curve' became an ongoing nightmare of continuing lockdowns with periods in between of severe restrictions in pursuit of destroying independent incomes and had nothing to do with protecting health about which the Cult gives not a shit. Why wouldn't Ferguson be pushing a vaccine 'solution' when he's owned by vaccine-obsessive Gates who makes a fortune from them and

when Ferguson heads the Vaccine Impact Modelling Consortium at Imperial College funded by the Gates Foundation and GAVI, the 'vaccine alliance', created by Gates as his personal vaccine promotion operation? To compound the human catastrophe that Ferguson's 'models' did so much to create he was later exposed for breaking his own lockdown rules by having sexual liaisons with his married girlfriend Antonia Staats at his home while she was living at another location with her husband and children. Staats was a 'climate' activist and senior campaigner at the Soros-funded Avaaz which I wouldn't trust to tell me that grass is green. Ferguson had to resign as a government advisor over this hypocrisy in May, 2020, but after a period of quiet he was back being quoted by the ridiculous media on the need for more lockdowns and a vaccine rollout. Other government-advising 'scientists' from Imperial College' held the fort in his absence and said lockdown could be indefinite until a vaccine was found. The Cult script was being sung by the payrolled choir. I said there was no intention of going back to 'normal' when the 'vaccine' came because the 'vaccine' is part of a very different agenda that I will discuss in Human 2.0. Why would the Cult want to let the world go back to normal when destroying that normal forever was the whole point of what was happening? House arrest, closing businesses and schools through lockdown, (un)social distancing and masks all followed the Ferguson fantasy models. Again as I predicted (these people are so predictable) when the 'vaccine' arrived we were told that house arrest, lockdown, (un)social distancing and masks would still have to continue. I will deal with the masks in the next chapter because they are of fundamental importance.

Where's the 'pandemic'?

Any mildly in-depth assessment of the figures revealed what was really going on. Cult-funded and controlled organisations still have genuine people working within them such is the number involved. So it is with Genevieve Briand, assistant program director of the Applied Economics master's degree program at Johns Hopkins

University. She analysed the impact that 'Covid-19' had on deaths from *all* causes in the United States using official data from the CDC for the period from early February to early September, 2020. She found that allegedly 'Covid' *related*-deaths exceeded those from heart disease which she found strange with heart disease always the biggest cause of fatalities. Her research became even more significant when she noted the sudden decline in 2020 of *all* non-'Covid' deaths: 'This trend is completely contrary to the pattern observed in all previous years … the total decrease in deaths by other causes almost exactly equals the increase in deaths by Covid-19.' This was such a game, set and match in terms of what was happening that Johns Hopkins University deleted the article on the grounds that it 'was being used to support false and dangerous inaccuracies about the impact of the pandemic'. No – because it exposed the scam from official CDC figures and this was confirmed when those figures were published in January, 2021. Here we can see the effect of people dying from heart attacks, cancer, road accidents and gunshot wounds – *anything* – having 'Covid-19' on the death certificate along with those diagnosed from 'symptoms' who had even not tested positive with a test not testing for the 'virus'. I am not kidding with the gunshot wounds, by the way. Brenda Bock, coroner in Grand County, Colorado, revealed that two gunshot victims tested positive for the 'virus' within the previous 30 days and were therefore classified as 'Covid deaths'. Bock said: 'These two people had tested positive for Covid, but that's not what killed them. A gunshot wound is what killed them.' She said she had not even finished her investigation when the state listed the gunshot victims as deaths due to the 'virus'. The death and case figures for 'Covid-19' are an absolute joke and yet they are repeated like parrots by the media, politicians and alleged medical 'experts'. The official Cult narrative is the only show in town.

Genevieve Briand found that deaths from all causes were not exceptional in 2020 compared with previous years and a Spanish magazine published figures that said the same about Spain which was a 'Covid' propaganda hotspot at one point. *Discovery Salud*, a

health and medicine magazine, quoted government figures which showed how 17,000 *fewer* people died in Spain in 2020 than in 2019 and more than 26,000 fewer than in 2018. The age-standardised mortality rate for England and Wales when age distribution is taken into account was significantly lower in 2020 than the 1970s, 80s and 90s, and was only the ninth highest since 2000. Where is the 'pandemic'?

Post mortems and autopsies virtually disappeared for 'Covid' deaths amid claims that 'virus-infected' bodily fluids posed a risk to those carrying out the autopsy. This was rejected by renowned German pathologist and forensic doctor Klaus Püschel who said that he and his staff had by then done 150 autopsies on 'Covid' patients with no problems at all. He said they were needed to know why some 'Covid' patients suffered blood clots and not severe respiratory infections. The 'virus' is, after all, called SARS or 'severe acute respiratory syndrome'. I highlighted in the spring of 2020 this phenomenon and quoted New York intensive care doctor Cameron Kyle-Sidell who posted a soon deleted YouTube video to say that they had been told to prepare to treat an infectious disease called 'Covid-19', but that was not what they were dealing with. Instead he likened the lung condition of the most severely ill patients to what you would expect with cabin depressurisation in a plane at 30,000 feet or someone dropped on the top of Everest without oxygen or acclimatisation. I have never said this is not happening to a small minority of alleged 'Covid' patients – I am saying this is not caused by a phantom 'contagious virus'. Indeed Kyle-Sidell said that 'Covid-19' was not the disease they were told was coming their way. 'We are operating under a medical paradigm that is untrue,' he said, and he believed they were treating the wrong disease: 'These people are being slowly starved of oxygen.' Patients would take off their oxygen masks in a state of fear and stress and while they were blue in the face on the brink of death. They did not look like patients dying of pneumonia. You can see why they don't want autopsies when their virus doesn't exist and there is another condition in some people that they don't wish to be uncovered. I should add here that

the 5G system of millimetre waves was being rapidly introduced around the world in 2020 and even more so now as they fire 5G at the Earth from satellites. At 60 gigahertz within the 5G range that frequency interacts with the oxygen molecule and stops people breathing in sufficient oxygen to be absorbed into the bloodstream. They are installing 5G in schools and hospitals. The world is not mad or anything. 5G can cause major changes to the lungs and blood as I detail in *The Answer* and these consequences are labelled 'Covid-19', the alleged symptoms of which can be caused by 5G and other electromagnetic frequencies as cells respond to radiation poisoning.

The 'Covid death' scam

Dr Scott Jensen, a Minnesota state senator and medical doctor, exposed 'Covid' Medicare payment incentives to hospitals and death certificate manipulation. He said he was sent a seven-page document by the US Department of Health 'coaching' him on how to fill out death certificates which had never happened before. The document said that he didn't need to have a laboratory test for 'Covid-19' to put that on the death certificate and that shocked him when death certificates are supposed to be about facts. Jensen described how doctors had been 'encouraged, if not pressured' to make a diagnosis of 'Covid-19' if they thought it was probable or *'presumed*'. No positive test was necessary – not that this would have mattered anyway. He said doctors were told to diagnose 'Covid' by symptoms when these were the same as colds, allergies, other respiratory problems, and certainly with influenza which 'disappeared' in the 'Covid' era. A common sniffle was enough to get the dreaded verdict. Ontario authorities decreed that a single care home resident with *one* symptom from a long list must lead to the isolation of the entire home. Other courageous doctors like Jensen made the same point about death figure manipulation and how deaths by other causes were falling while 'Covid-19 deaths' were rising at the same rate due to re-diagnosis. Their videos rarely survive long on YouTube with its Cult-supporting algorithms courtesy of CEO Susan Wojcicki and her bosses at Google. Figure-tampering was so glaring

and ubiquitous that even officials were letting it slip or outright saying it. UK chief scientific adviser Patrick Vallance said on one occasion that 'Covid' on the death certificate doesn't mean 'Covid' was the cause of death (so why the hell is it there?) and we had the rare sight of a BBC reporter telling the truth when she said: 'Someone could be successfully treated for Covid, in say April, discharged, and then in June, get run over by a bus and die … That person would still be counted as a Covid death in England.' Yet the BBC and the rest of the world media went on repeating the case and death figures as if they were real. Illinois Public Health Director Dr Ngozi Ezike revealed the deceit while her bosses must have been clenching their buttocks:

> If you were in a hospice and given a few weeks to live and you were then found to have Covid that would be counted as a Covid death. [There might be] a clear alternate cause, but it is still listed as a Covid death. So everyone listed as a Covid death doesn't mean that was the cause of the death, but that they had Covid at the time of death.

Yes, a 'Covid virus' never shown to exist and tested for with a test not testing for the 'virus'. In the first period of the pandemic hoax through the spring of 2020 the process began of designating almost everything a 'Covid' death and this has continued ever since. I sat in a restaurant one night listening to a loud conversation on the next table where a family was discussing in bewilderment how a relative who had no symptoms of 'Covid', and had died of a long-term problem, could have been diagnosed a death by the 'virus'. I could understand their bewilderment. If they read this book they will know why this medical fraud has been perpetrated the world over.

Some media truth shock

The media ignored the evidence of death certificate fraud until eventually one columnist did speak out when she saw it first-hand. Bel Mooney is a long-time national newspaper journalist in Britain currently working for the *Daily Mail*. Her article on February 19th, 2021, carried this headline: 'My dad Ted passed three Covid tests

and died of a chronic illness yet he's officially one of Britain's 120,000 victims of the virus and is far from alone ... so how many more are there?' She told how her 99-year-old father was in a care home with a long-standing chronic obstructive pulmonary disease and vascular dementia. Maybe, but he was still aware enough to tell her from the start that there was no 'virus' and he refused the 'vaccine' for that reason. His death was not unexpected given his chronic health problems and Mooney said she was shocked to find that 'Covid-19' was declared the cause of death on his death certificate. She said this was a 'bizarre and unacceptable untruth' for a man with long-time health problems who had tested negative twice at the home for the 'virus'. I was also shocked by this story although not by what she said. I had been highlighting the death certificate manipulation for ten months. It was the confirmation that a professional full-time journalist only realised this was going on when it affected her directly and neither did she know that whether her dad tested positive or negative was irrelevant with the test not testing for the 'virus'. Where had she been? She said she did not believe in 'conspiracy theories' without knowing I'm sure that this and 'conspiracy theorists' were terms put into widespread circulation by the CIA in the 1960s to discredit those who did not accept the ridiculous official story of the Kennedy assassination. A blanket statement of 'I don't believe in conspiracy theories' is always bizarre. The dictionary definition of the term alone means the world is drowning in conspiracies. What she said was even more daft when her dad had just been affected by the 'Covid' conspiracy. Why else does she think that 'Covid-19' was going on the death certificates of people who died of something else?

To be fair once she saw from personal experience what was happening she didn't mince words. Mooney was called by the care home on the morning of February 9th to be told her father had died in his sleep. When she asked for the official cause of death what came back was 'Covid-19'. Mooney challenged this and was told there had been deaths from Covid on the dementia floor (confirmed by a test not testing for the 'virus') so they considered it 'reasonable

to assume'. 'But doctor,' Mooney rightly protested, 'an assumption isn't a diagnosis.' She said she didn't blame the perfectly decent and sympathetic doctor – 'he was just doing his job'. Sorry, but that's *bullshit*. He wasn't doing his job at all. He was putting a false cause of death on the death certificate and that is a criminal offence for which he should be brought to account and the same with the millions of doctors worldwide who have done the same. They were not doing their job they were following orders and that must not wash at new Nuremberg trials any more than it did at the first ones. Mooney's doctor was 'assuming' (presuming) as he was told to, but 'just following orders' makes no difference to his actions. A doctor's job is to serve the patient and the truth, not follow orders, but that's what they have done all over the world and played a central part in making the 'Covid' hoax possible with all its catastrophic consequences for humanity. Shame on them and they must answer for their actions. Mooney said her disquiet worsened when she registered her father's death by telephone and was told by the registrar there had been very many other cases like hers where 'the deceased' had not tested positive for 'Covid' yet it was recorded as the cause of death. The test may not matter, but those involved at their level *think* it matters and it shows a callous disregard for accurate diagnosis. The pressure to do this is coming from the top of the national 'health' pyramids which in turn obey the World Health Organization which obeys Gates and the Cult. Mooney said the registrar agreed that this must distort the national figures adding that 'the strangest thing is that every winter we record countless deaths from flu, and this winter there have been none. Not one!' She asked if the registrar thought deaths from flu were being misdiagnosed and lumped together with 'Covid' deaths. The answer was a 'puzzled yes'. Mooney said that the funeral director said the same about 'Covid' deaths which had nothing to do with 'Covid'. They had lost count of the number of families upset by this and other funeral companies in different countries have had the same experience. Mooney wrote:

The nightly shroud-waving and shocking close-ups of pain imposed on us by the TV news bewildered and terrified the population into eager compliance with lockdowns. We were invited to 'save the NHS' and to grieve for strangers – the real-life loved ones behind those shocking death counts. Why would the public imagine what I now fear, namely that the way Covid-19 death statistics are compiled might make the numbers seem greater than they are?

Oh, just a little bit – like 100 percent.

Do the maths

Mooney asked why a country would wish to skew its mortality figures by wrongly certifying deaths? What had been going on? Well, if you don't believe in conspiracies you will never find the answer which is that *it's a conspiracy*. She did, however, describe what she had discovered as a 'national scandal'. In reality it's a global scandal and happening everywhere. Pillars of this conspiracy were all put into place before the button was pressed with the Drosten PCR protocol and high amplifications to produce the cases and death certificate changes to secure illusory 'Covid' deaths. Mooney notes that normally two doctors were needed to certify a death, with one having to know the patient, and how the rules were changed in the spring of 2020 to allow one doctor to do this. In the same period 'Covid deaths' were decreed to be all cases where Covid-19 was put on the death certificate even without a positive test or any symptoms. Mooney asked: 'How many of the 30,851 (as of January 15) care home resident deaths with Covid-19 on the certificate (32.4 per cent of all deaths so far) were based on an assumption, like that of my father? And what has that done to our national psyche?'All of them is the answer to the first question and it has devastated and dismantled the national psyche, actually the global psyche, on a colossal scale. In the UK case and death data is compiled by organisations like Public Health England (PHE) and the Office for National Statistics (ONS). Mooney highlights the insane policy of counting a death from any cause as 'Covid-19' if this happens within 28 days of a positive test (with a test not testing for the 'virus') and she points out that ONS statistics reflect deaths 'involving Covid' 'or due to Covid' which meant in practice any

death where ‘Covid-19’ was mentioned on the death certificate. She described the consequences of this fraud:

Most people will accept the narrative they are fed, so panicky governments here and in Europe witnessed the harsh measures enacted in totalitarian China and jumped into lockdown. Headlines about Covid deaths tolled like the knell that would bring doomsday to us all. Fear stalked our empty streets. Politicians parroted the frankly ridiculous aim of ‘zero Covid’ and shut down the economy, while most British people agreed that lockdown was essential and (astonishingly to me, as a patriotic Brit) even wanted more restrictions.

For what? Lies on death certificates? Never mind the grim toll of lives ruined, suicides, schools closed, rising inequality, depression, cancelled hospital treatments, cancer patients in a torture of waiting, poverty, economic devastation, loneliness, families kept apart, and so on. How many lives have been lost as a direct result of lockdown?

She said that we could join in a national chorus of shock and horror at reaching the 120,000 death toll which was surely certain to have been totally skewed all along, but what about the human cost of lockdown justified by these ‘death figures’? *The British Medical Journal* had reported a 1,493 percent increase in cases of children taken to Great Ormond Street Hospital with abusive head injuries alone and then there was the effect on families:

Perhaps the most shocking thing about all this is that families have been kept apart – and obeyed the most irrational, changing rules at the whim of government – because they believed in the statistics. They succumbed to fear, which his generation rejected in that war fought for freedom. Dad (God rest his soul) would be angry. And so am I.

Another theme to watch is that in the winter months when there are more deaths from all causes they focus on ‘Covid’ deaths and in the summer when the British Lung Foundation says respiratory disease plummets by 80 percent they rage on about ‘cases’. Either way fascism on population is always the answer.

Nazi eugenics in the 21st century

Elderly people in care homes have been isolated from their families month after lonely month with no contact with relatives and grandchildren who were banned from seeing them. We were told

that lockdown fascism was to 'protect the vulnerable' like elderly people. At the same time Do Not Resuscitate (DNR) orders were placed on their medical files so that if they needed resuscitation it wasn't done and 'Covid-19' went on their death certificates. Old people were not being 'protected' they were being culled – murdered in truth. DNR orders were being decreed for disabled and young people with learning difficulties or psychological problems. The UK Care Quality Commission, a non-departmental body of the Department of Health and Social Care, found that 34 percent of those working in health and social care were pressured into placing 'do not attempt cardiopulmonary resuscitation' orders on 'Covid' patients who suffered from disabilities and learning difficulties without involving the patient or their families in the decision. UK judges ruled that an elderly woman with dementia should have the DNA-manipulating 'Covid vaccine' against her son's wishes and that a man with severe learning difficulties should have the jab despite his family's objections. Never mind that many had already died. The judiciary always supports doctors and government in fascist dictatorships. They wouldn't dare do otherwise. A horrific video was posted showing fascist officers from Los Angeles police forcibly giving the 'Covid' shot to women with special needs who were screaming that they didn't want it. The same fascists are seen giving the jab to a sleeping elderly woman in a care home. This is straight out of the Nazi playbook. Hitler's Nazis committed mass murder of the mentally ill and physically disabled throughout Germany and occupied territories in the programme that became known as Aktion T4, or just T4. Sabbatian-controlled Hitler and his grotesque crazies set out to kill those they considered useless and unnecessary. The Reich Committee for the Scientific Registering of Hereditary and Congenital Illnesses registered the births of babies identified by physicians to have 'defects'. By 1941 alone more than 5,000 children were murdered by the state and it is estimated that in total the number of innocent people killed in Aktion T4 was between 275,000 and 300,000. Parents were told their children had been sent away for 'special treatment' never to return. It is rather pathetic to see claims about plans for new extermination camps being dismissed today

when the same force behind current events did precisely that 80 years ago. Margaret Sanger was a Cult operative who used 'birth control' to sanitise her programme of eugenics. Organisations she founded became what is now Planned Parenthood. Sanger proposed that 'the whole dysgenic population would have its choice of segregation or sterilization'. These included epileptics, 'feeble-minded', and prostitutes. Sanger opposed charity because it perpetuated 'human waste'. She reveals the Cult mentality and if anyone thinks that extermination camps are a 'conspiracy theory' their naivety is touching if breathtakingly stupid.

If you don't believe that doctors can act with callous disregard for their patients it is worth considering that doctors and medical staff agreed to put government-decreed DNR orders on medical files and do nothing when resuscitation is called for. I don't know what you call such people in your house. In mine they are Nazis from the Josef Mengele School of Medicine. Phenomenal numbers of old people have died worldwide from the effects of lockdown, depression, lack of treatment, the 'vaccine' (more later) and losing the will to live. A common response at the start of the manufactured pandemic was to remove old people from hospital beds and transfer them to nursing homes. The decision would result in a mass cull of elderly people in those homes through lack of treatment – *not* 'Covid'. Care home whistleblowers have told how once the 'Covid' era began doctors would not come to their homes to treat patients and they were begging for drugs like antibiotics that often never came. The most infamous example was ordered by New York governor Andrew Cuomo, brother of a moronic CNN host, who amazingly was given an Emmy Award for his handling of the 'Covid crisis' by the ridiculous Wokers that hand them out. Just how ridiculous could be seen in February, 2021, when a Department of Justice and FBI investigation began into how thousands of old people in New York died in nursing homes after being discharged from hospital to make way for 'Covid' patients on Cuomo's say-so – and how he and his staff covered up these facts. This couldn't have happened to a nicer psychopath. Even then there was a 'Covid' spin. Reports said that

thousands of old people who tested positive for 'Covid' in hospital were transferred to nursing homes to both die of 'Covid' and transmit it to others. No – they were in hospital because they were ill and the fact that they tested positive with a test not testing for the 'virus' is irrelevant. They were ill often with respiratory diseases ubiquitous in old people near the end of their lives. Their transfer out of hospital meant that their treatment stopped and many would go on to die.

They're old. Who gives a damn?

I have exposed in the books for decades the Cult plan to cull the world's old people and even to introduce at some point what they call a 'demise pill' which at a certain age everyone would take and be out of here by law. In March, 2021, Spain legalised euthanasia and assisted suicide following the Netherlands, Belgium, Luxembourg and Canada on the Tiptoe to the demise pill. Treatment of old people by many 'care' homes has been a disgrace in the 'Covid' era. There are many, many, caring staff – I know some. There have, however, been legions of stories about callous treatment of old people and their families. Police were called when families came to take their loved ones home in the light of isolation that was killing them. They became prisoners of the state. Care home residents in insane, fascist Ontario, Canada, were not allowed to leave their *room* once the 'Covid' hoax began. UK staff have even wheeled elderly people away from windows where family members were talking with them. Oriana Criscuolo from Stockport in the English North West dropped off some things for her 80-year-old father who has Parkinson's disease and dementia and she wanted to wave to him through a ground-floor window. She was told that was 'illegal'. When she went anyway they closed the curtains in the middle of the day. Oriana said:

> It's just unbelievable. I cannot understand how care home staff – people who are being paid to care – have become so uncaring. Their behaviour is inhumane and cruel. It's beyond belief.

She was right and this was not a one-off. What a way to end your life in such loveless circumstances. UK registered nurse Nicky Millen, a proper old school nurse for 40 years, said that when she started her career care was based on dignity, choice, compassion and empathy. Now she said 'the things that are important to me have gone out of the window.' She was appalled that people were dying without their loved ones and saying goodbye on iPads. Nicky described how a distressed 89-year-old lady stroked her face and asked her 'how many paracetamol would it take to finish me off'. Life was no longer worth living while not seeing her family. Nicky said she was humiliated in front of the ward staff and patients for letting the lady stroke her face and giving her a cuddle. Such is the dehumanisation that the 'Covid' hoax has brought to the surface. Nicky worked in care homes where patients told her they were being held prisoner. 'I want to live until I die', one said to her. 'I had a lady in tears because she hadn't seen her great-grandson.' Nicky was compassionate old school meeting psychopathic New Normal. She also said she had worked on a 'Covid' ward with no 'Covid' patients. Jewish writer Shai Held wrote an article in March, 2020, which was headlined 'The Staggering, Heartless Cruelty Toward the Elderly'. What he described was happening from the earliest days of lockdown. He said 'the elderly' were considered a group and not unique individuals (the way of the Woke). Shai Held said:

> Notice how the all-too-familiar rhetoric of dehumanization works: 'The elderly' are bunched together as a faceless mass, all of them considered culprits and thus effectively deserving of the suffering the pandemic will inflict upon them. Lost entirely is the fact that the elderly are individual human beings, each with a distinctive face and voice, each with hopes and dreams, memories and regrets, friendships and marriages, loves lost and loves sustained.

'The elderly' have become another dehumanised group for which anything goes and for many that has resulted in cold disregard for their rights and their life. The distinctive face that Held talks about is designed to be deleted by masks until everyone is part of a faceless mass.

‘War-zone’ hospitals myth

Again and again medical professionals have told me what was really going on and how hospitals ‘overrun like war zones’ according to the media were virtually empty. The mantra from medical whistleblowers was please don’t use my name or my career is over. Citizen journalists around the world sneaked into hospitals to film evidence exposing the ‘war-zone’ lie. They really *were* largely empty with closed wards and operating theatres. I met a hospital worker in my town on the Isle of Wight during the first lockdown in 2020 who said the only island hospital had never been so quiet. Lockdown was justified by the psychopaths to stop hospitals being overrun. At the same time that the island hospital was near-empty the military arrived here to provide *extra beds*. It was all propaganda to ramp up the fear to ensure compliance with fascism as were never-used temporary hospitals with thousands of beds known as Nightingales and never-used make-shift mortuaries opened by the criminal UK government. A man who helped to install those extra island beds attributed to the army said they were never used and the hospital was empty. Doctors and nurses ‘stood around talking or on their phones, wandering down to us to see what we were doing’. There were no masks or social distancing. He accused the useless local island paper, the *County Press*, of ‘pumping the fear as if our hospital was overrun and we only have one so it should have been’. He described ambulances parked up with crews outside in deck chairs. When his brother called an ambulance he was told there was a two-hour backlog which he called ‘bullshit’. An old lady on the island fell ‘and was in a bad way’, but a caller who rang for an ambulance was told the situation wasn’t urgent enough. Ambulance stations were working under capacity while people would hear ambulances with sirens blaring driving through the streets. When those living near the stations realised what was going on they would follow them as they left, circulated around an urban area with the sirens going, and then came back without stopping. All this was to increase levels of fear and the same goes for the ‘ventilator shortage crisis’ that cost tens of millions for hastily produced ventilators never to be used.

Ambulance crews that agreed to be exploited in this way for fear propaganda might find themselves a mirror. I wish them well with that. Empty hospitals were the obvious consequence of treatment and diagnoses of non-'Covid' conditions cancelled and those involved handed a death sentence. People have been dying at home from undiagnosed and untreated cancer, heart disease and other life-threatening conditions to allow empty hospitals to deal with a 'pandemic' that wasn't happening.

Death of the innocent

'War-zones' have been laying off nursing staff, even doctors where they can. There was no work for them. Lockdown was justified by saving lives and protecting the vulnerable they were actually killing with DNR orders and preventing empty hospitals being 'overrun'. In Britain the mantra of stay at home to 'save the NHS' was everywhere and across the world the same story was being sold when it was all lies. Two California doctors, Dan Erickson and Artin Massihi at Accelerated Urgent Care in Bakersfield, held a news conference in April, 2020, to say that intensive care units in California were 'empty, essentially', with hospitals shutting floors, not treating patients and laying off doctors. The California health system was working at minimum capacity 'getting rid of doctors because we just don't have the volume'. They said that people with conditions such as heart disease and cancer were not coming to hospital out of fear of 'Covid-19'. Their video was deleted by Susan Wojcicki's Cult-owned YouTube after reaching five million views. Florida governor Ron Desantis, who rejected the severe lockdowns of other states and is being targeted for doing so, said that in March, 2020, every US governor was given models claiming they would run out of hospital beds in days. That was never going to happen and the 'modellers' knew it. Deceit can be found at every level of the system. Urgent children's operations were cancelled including fracture repairs and biopsies to spot cancer. Eric Nicholls, a consultant paediatrician, said 'this is obviously concerning and we need to return to normal operating and to increase capacity as soon as possible'. Psychopaths

in power were rather less concerned *because* they are psychopaths. Deletion of urgent care and diagnosis has been happening all over the world and how many kids and others have died as a result of the actions of these cold and heartless lunatics dictating 'health' policy? The number must be stratospheric. Richard Sullivan, professor of cancer and global health at King's College London, said people feared 'Covid' more than cancer such was the campaign of fear. 'Years of lost life will be quite dramatic', Sullivan said, with 'a huge amount of avoidable mortality'. Sarah Woolnough, executive director for policy at Cancer Research UK, said there had been a 75 percent drop in urgent referrals to hospitals by family doctors of people with suspected cancer. Sullivan said that 'a lot of services have had to scale back – we've seen a dramatic decrease in the amount of elective cancer surgery'. Lockdown deaths worldwide has been absolutely fantastic with the *New York Post* reporting how data confirmed that 'lockdowns end more lives than they save':

> There was a sharp decline in visits to emergency rooms and an increase in fatal heart attacks because patients didn't receive prompt treatment. Many fewer people were screened for cancer. Social isolation contributed to excess deaths from dementia and Alzheimer's.
>
> Researchers predicted that the social and economic upheaval would lead to tens of thousands of "deaths of despair" from drug overdoses, alcoholism and suicide. As unemployment surged and mental-health and substance-abuse treatment programs were interrupted, the reported levels of anxiety, depression and suicidal thoughts increased dramatically, as did alcohol sales and fatal drug overdoses.

This has been happening while nurses and other staff had so much time on their hands in the 'war-zones' that Tic-Tok dancing videos began appearing across the Internet with medical staff dancing around in empty wards and corridors as people died at home from causes that would normally have been treated in hospital.

Mentions in dispatches

One brave and truth-committed whistleblower was Louise Hampton, a call handler with the UK NHS who made a viral Internet video saying she had done 'fuck all' during the 'pandemic'

which was 'a load of bollocks'. She said that 'Covid-19' was rebranded flu and of course she lost her job. This is what happens in the medical and endless other professions now when you tell the truth. Louise filmed inside 'war-zone' accident and emergency departments to show they were empty and I mean *empty* as in no one there. The mainstream media could have done the same and blown the gaff on the whole conspiracy. They haven't to their eternal shame. Not that most 'journalists' seem capable of manifesting shame as with the psychopaths they slavishly repeat without question. The relative few who were admitted with serious health problems were left to die alone with no loved ones allowed to see them because of 'Covid' rules and they included kids dying without the comfort of mum and dad at their bedside while the evil behind this couldn't give a damn. It was all good fun to them. A Scottish NHS staff nurse publicly quit in the spring of 2021 saying: 'I can no longer be part of the lies and the corruption by the government.' She said hospitals 'aren't full, the beds aren't full, beds have been shut, wards have been shut'. Hospitals were never busy throughout 'Covid'. The staff nurse said that Nicola Sturgeon, tragically the leader of the Scottish government, was on television saying save the hospitals and the NHS – 'but the beds are empty' and 'we've not seen flu, we always see flu every year'. She wrote to government and spoke with her union Unison (the unions are Cult-compromised and *useless*, but nothing changed. Many of her colleagues were scared of losing their jobs if they spoke out as they wanted to. She said nursing staff were being affected by wearing masks all day and 'my head is splitting every shift from wearing a mask'. The NHS is part of the fascist tyranny and must be dismantled so we can start again with human beings in charge. (Ironically, hospitals were reported to be busier again when official 'Covid' cases *fell* in spring/summer of 2021 and many other conditions required treatment at the same time as *the fake vaccine rollout*.)

I will cover the 'Covid vaccine' scam in detail later, but it is another indicator of the sickening disregard for human life that I am highlighting here. The DNA-manipulating concoctions do not fulfil

the definition of a 'vaccine', have never been used on humans before and were given only emergency approval because trials were not completed and they continued using the unknowing public. The result was what a NHS senior nurse with responsibility for 'vaccine' procedure said was 'genocide'. She said the 'vaccines' were not 'vaccines'. They had not been shown to be safe and claims about their effectiveness by drug companies were 'poetic licence'. She described what was happening as a 'horrid act of human annihilation'. The nurse said that management had instigated a policy of not providing a Patient Information Leaflet (PIL) before people were 'vaccinated' even though health care professionals are supposed to do this according to protocol. Patients should also be told that they are taking part in an ongoing clinical trial. Her challenges to what is happening had seen her excluded from meetings and ridiculed in others. She said she was told to 'watch my step … or I would find myself surplus to requirements'. The nurse, who spoke anonymously in fear of her career, said she asked her NHS manager why he/she was content with taking part in genocide against those having the 'vaccines'. The reply was that everyone had to play their part and to 'put up, shut up, and get it done'. Government was 'leaning heavily' on NHS management which was clearly leaning heavily on staff. This is how the global 'medical' hierarchy operates and it starts with the Cult and its World Health Organization.

She told the story of a doctor who had the Pfizer jab and when questioned had no idea what was in it. The doctor had never read the literature. We have to stop treating doctors as intellectual giants when so many are moral and medical pygmies. The doctor did not even know that the 'vaccines' were not fully approved or that their trials were ongoing. They were, however, asking their patients if they minded taking part in follow-ups for research purposes – yes, the *ongoing clinical trial*. The nurse said the doctor's ignorance was not rare and she had spoken to a hospital consultant who had the jab without any idea of the background or that the 'trials' had not been completed. Nurses and pharmacists had shown the same ignorance.

'My NHS colleagues have forsaken their duty of care, broken their code of conduct – Hippocratic Oath – and have been brainwashed just the same as the majority of the UK public through propaganda …' She said she had not been able to recruit a single NHS colleague, doctor, nurse or pharmacist to stand with her and speak out. Her union had refused to help. She said that if the genocide came to light she would not hesitate to give evidence at a Nuremberg-type trial against those in power who could have affected the outcomes but didn't.

And all for what?

To put the nonsense into perspective let's say the 'virus' does exist and let's go completely crazy and accept that the official manipulated figures for cases and deaths are accurate. *Even then* a study by Stanford University epidemiologist Dr John Ioannidis published on the World Health Organization website produced an average infection to fatality rate of … *0.23 percent!* Ioannidis said: 'If one could sample equally from all locations globally, the median infection fatality rate might even be substantially lower than the 0.23% observed in my analysis.' For healthy people under 70 it was … *0.05 percent*! This compares with the 3.4 percent claimed by the Cult-owned World Health Organization when the hoax was first played and maximum fear needed to be generated. An updated Stanford study in April, 2021, put the 'infection' to 'fatality' rate at just 0.15 percent. Another team of scientists led by Megan O'Driscoll and Henrik Salje studied data from 45 countries and published their findings on the Nature website. For children and young people the figure is so small it virtually does not register although authorities will be hyping dangers to the young when they introduce DNA-manipulating 'vaccines' for children. The O'Driscoll study produced an average infection-fatality figure of 0.003 for children from birth to four; 0.001 for 5 to 14; 0.003 for 15 to 19; and it was still only 0.456 up to 64. To claim that children must be 'vaccinated' to protect them from 'Covid' is an obvious lie and so there must be another reason and there is. What's more the average age of a 'Covid' death is akin

to the average age that people die in general. The average age of death in England is about 80 for men and 83 for women. The average age of death from alleged 'Covid' is between 82 and 83. California doctors, Dan Erickson and Artin Massihi, said at their April media conference that projection models of millions of deaths had been 'woefully inaccurate'. They produced detailed figures showing that Californians had a 0.03 chance of dying from 'Covid' based on the number of people who tested positive (with a test not testing for the 'virus'). Erickson said there was a 0.1 percent chance of dying from 'Covid' in the *state* of New York, not just the city, and a 0.05 percent chance in Spain, a centre of 'Covid-19' hysteria at one stage. The Stanford studies supported the doctors' data with fatality rate estimates of 0.23 and 0.15 percent. How close are these figures to my estimate of *zero*? Death-rate figures claimed by the World Health Organization at the start of the hoax were some 15 times higher. The California doctors said there was no justification for lockdowns and the economic devastation they caused. Everything they had ever learned about quarantine was that you quarantine the *sick* and not the healthy. They had never seen this before and it made no medical sense.

Why in the in the light of all this would governments and medical systems the world over say that billions must go under house arrest; lose their livelihood; in many cases lose their mind, their health and their life; force people to wear masks dangerous to health and psychology; make human interaction and even family interaction a criminal offence; ban travel; close restaurants, bars, watching live sport, concerts, theatre, and any activity involving human togetherness and discourse; and closing schools to isolate children from their friends and cause many to commit suicide in acts of hopelessness and despair? The California doctors said lockdown consequences included increased child abuse, partner abuse, alcoholism, depression, and other impacts they were seeing every day. Who would do that to the entire human race if not mentally-ill psychopaths of almost unimaginable extremes like Bill Gates? We must face the reality of what we are dealing with and come out of

denial. Fascism and tyranny are made possible only by the target population submitting and acquiescing to fascism and tyranny. The whole of human history shows that to be true. Most people naively and unquestioning believed what they were told about a 'deadly virus' and meekly and weakly submitted to house arrest. Those who didn't believe it – at least in total – still submitted in fear of the consequences of not doing so. For the rest who wouldn't submit draconian fines have been imposed, brutal policing by psychopaths *for* psychopaths, and condemnation from the meek and weak who condemn the Pushbackers on behalf of the very force that has them, too, in its gunsights. 'Pathetic' does not even begin to suffice. Britain's brainless 'Health' Secretary Matt Hancock warned anyone lying to border officials about returning from a list of 'hotspot' countries could face a jail sentence of up to ten years which is more than for racially-aggravated assault, incest and attempting to have sex with a child under 13. Hancock is a lunatic, but he has the state apparatus behind him in a Cult-led chain reaction and the same with UK 'Vaccine Minister' Nadhim Zahawi, a prominent member of the mega-Cult secret society, Le Cercle, which featured in my earlier books. The Cult enforces its will on governments and medical systems; government and medical systems enforce their will on business and police; business enforces its will on staff who enforce it on customers; police enforce the will of the Cult on the population and play their essential part in creating a world of fascist control that their own children and grandchildren will have to live in their entire lives. It is a hierarchical pyramid of imposition and acquiescence and, yes indeedy, of clinical insanity.

Does anyone bright enough to read this book have to ask what the answer is? I think not, but I will reveal it anyway in the fewest of syllables: Tell the psychos and their moronic lackeys to fuck off and let's get on with our lives. We are many – They are few.

CHAPTER SEVEN

War on your mind

One believes things because one has been conditioned to believe them
Aldous Huxley, *Brave New World*

I have described the 'Covid' hoax as a 'Psyop' and that is true in every sense and on every level in accordance with the definition of that term which is psychological warfare. Break down the 'Covid pandemic' to the foundation themes and it is psychological warfare on the human individual and collective mind.

The same can be said for the entire human belief system involving every subject you can imagine. Huxley was right in his contention that people believe what they are conditioned to believe and this comes from the repetition throughout their lives of the same falsehoods. They spew from government, corporations, media and endless streams of 'experts' telling you what the Cult wants you to believe and often believing it themselves (although *far* from always). 'Experts' are rewarded with 'prestigious' jobs and titles and as agents of perceptual programming with regular access to the media. The Cult has to control the narrative – control *information* – or they lose control of the vital, crucial, without-which-they-cannot-prevail public perception of reality. The foundation of that control today is the Internet made possible by the Defense Advanced Research Projects Agency (DARPA), the incredibly sinister technological arm of the Pentagon. The Internet is the result of military technology.

DARPA openly brags about establishing the Internet which has been a long-term project to lasso the minds of the global population. I have said for decades the plan is to control information to such an extreme that eventually no one would see or hear anything that the Cult does not approve. We are closing in on that end with ferocious censorship since the 'Covid' hoax began and in my case it started back in the 1990s in terms of books and speaking venues. I had to create my own publishing company in 1995 precisely because no one else would publish my books even then. I think they're all still running.

Cult Internet

To secure total control of information they needed the Internet in which pre-programmed algorithms can seek out 'unclean' content for deletion and even stop it being posted in the first place. The Cult had to dismantle print and non-Internet broadcast media to ensure the transfer of information to the appropriate-named 'Web' – a critical expression of the *Cult* web. We've seen the ever-quickening demise of traditional media and control of what is left by a tiny number of corporations operating worldwide. Independent journalism in the mainstream is already dead and never was that more obvious than since the turn of 2020. The Cult wants all information communicated via the Internet to globally censor and allow the plug to be pulled any time. Lockdowns and forced isolation has meant that communication between people has been through electronic means and no longer through face-to-face discourse and discussion. Cult psychopaths have targeted the bars, restaurants, sport, venues and meeting places in general for this reason. None of this is by chance and it's to stop people gathering in any kind of privacy or number while being able to track and monitor all Internet communications and block them as necessary. Even private messages between individuals have been censored by these fascists that control Cult fronts like Facebook, Twitter, Google and YouTube which are all officially run by Sabbatian place-people and from the background by higher-level Sabbatian place people.

Facebook, Google, Amazon and their like were seed-funded and supported into existence with money-no-object infusions of funds either directly or indirectly from DARPA and CIA technology arm In-Q-Tel. The Cult plays the long game and prepares very carefully for big plays like 'Covid'. Amazon is another front in the psychological war and pretty much controls the global market in book sales and increasingly publishing. Amazon's limitless funds have deleted fantastic numbers of independent publishers to seize global domination on the way to deciding which books can be sold and circulated and which cannot. Moves in that direction are already happening. Amazon's leading light Jeff Bezos is the grandson of Lawrence Preston Gise who worked with DARPA predecessor ARPA. Amazon has big connections to the CIA and the Pentagon. The plan I have long described went like this:

1. Employ military technology to establish the Internet.

2. Sell the Internet as a place where people can freely communicate without censorship and allow that to happen until the Net becomes the central and irreversible pillar of human society. If the Internet had been highly censored from the start many would have rejected it.

3. Fund and manipulate major corporations into being to control the circulation of information on your Internet using cover stories about geeks in garages to explain how they came about. Give them unlimited funds to expand rapidly with no need to make a profit for years while non-Cult companies who need to balance the books cannot compete. You know that in these circumstances your Googles, YouTubes, Facebooks and Amazons are going to secure near monopolies by either crushing or buying up the opposition.

4. Allow freedom of expression on both the Internet and communication platforms to draw people in until the Internet is the central and irreversible pillar of human society and your communication corporations have reached a stage of near monopoly domination.

5. Then unleash your always-planned frenzy of censorship on the basis of 'where else are you going to go?' and continue to expand that until nothing remains that the Cult does not want its human targets to see.

The process was timed to hit the 'Covid' hoax to ensure the best chance possible of controlling the narrative which they knew they had to do at all costs. They were, after all, about to unleash a 'deadly virus' that didn't really exist. If you do that in an environment of free-flowing information and opinion you would be dead in the

water before you could say Gates is a psychopath. The network was in place through which the Cult-created-and-owned World Health Organization could dictate the 'Covid' narrative and response policy slavishly supported by Cult-owned Internet communication giants and mainstream media while those telling a different story were censored. Google, YouTube, Facebook and Twitter openly announced that they would do this. What else would we expect from Cult-owned operations like Facebook which former executives have confirmed set out to make the platform more addictive than cigarettes and coldly manipulates emotions of its users to sow division between people and groups and scramble the minds of the young? If Zuckerberg lives out the rest of his life without going to jail for crimes against humanity, and most emphatically against the young, it will be a travesty of justice. Still, no matter, cause and effect will catch up with him eventually and the same with Sergey Brin and Larry Page at Google with its CEO Sundar Pichai who fix the Google search results to promote Cult narratives and hide the opposition. Put the same key words into Google and other search engines like DuckDuckGo and you will see how different results can be. Wikipedia is another intensely biased 'encyclopaedia' which skews its content to the Cult agenda. YouTube links to Wikipedia's version of 'Covid' and 'climate change' on video pages in which experts in their field offer a different opinion (even that is increasingly rare with Wojcicki censorship). Into this 'Covid' silence-them network must be added government media censors, sorry 'regulators', such as Ofcom in the UK which imposed tyrannical restrictions on British broadcasters that had the effect of banning me from ever appearing. Just to debate with me about my evidence and views on 'Covid' would mean breaking the fascistic impositions of Ofcom and its CEO career government bureaucrat Melanie Dawes. Gutless British broadcasters tremble at the very thought of fascist Ofcom.

Psychos behind 'Covid'

The reason for the 'Covid' catastrophe in all its facets and forms can be seen by whom and what is driving the policies worldwide in such a coordinated way. Decisions are not being made to protect health, but to target psychology. The dominant group guiding and 'advising' government policy are not medical professionals. They are psychologists and behavioural scientists. Every major country has its own version of this phenomenon and I'll use the British example to show how it works. In many ways the British version has been affecting the wider world in the form of the huge behaviour manipulation network in the UK which operates in other countries. The network involves private companies, government, intelligence and military. The Cabinet Office is at the centre of the government 'Covid' Psyop and part-owns, with 'innovation charity' Nesta, the Behavioural Insights Team (BIT) which claims to be independent of government but patently isn't. The BIT was established in 2010 and its job is to manipulate the psyche of the population to acquiesce to government demands and so much more. It is also known as the 'Nudge Unit', a name inspired by the 2009 book by two ultra-Zionists, Cass Sunstein and Richard Thaler, called *Nudge: Improving Decisions About Health, Wealth, and Happiness*. The book, as with the Behavioural Insights Team, seeks to 'nudge' behaviour (manipulate it) to make the public follow patterns of action and perception that suit those in authority (the Cult). Sunstein is so skilled at this that he advises the World Health Organization and the UK Behavioural Insights Team and was Administrator of the White House Office of Information and Regulatory Affairs in the Obama administration. Biden appointed him to the Department of Homeland Security – another ultra-Zionist in the fold to oversee new immigration laws which is another policy the Cult wants to control. Sunstein is desperate to silence anyone exposing conspiracies and co-authored a 2008 report on the subject in which suggestions were offered to ban 'conspiracy theorizing' or impose 'some kind of tax, financial or otherwise, on those who disseminate such theories'. I guess a psychiatrist's chair is out of the question?

Sunstein's mate Richard Thaler, an 'academic affiliate' of the UK Behavioural Insights Team, is a proponent of 'behavioural economics' which is defined as the study of 'the effects of psychological, cognitive, emotional, cultural and social factors on the decisions of individuals and institutions'. Study the effects so they can be manipulated to be what you want them to be. Other leading names in the development of behavioural economics are ultra-Zionists Daniel Kahneman and Robert J. Shiller and they, with Thaler, won the Nobel Memorial Prize in Economic Sciences for their work in this field. The Behavioural Insights Team is operating at the heart of the UK government and has expanded globally through partnerships with several universities including Harvard, Oxford, Cambridge, University College London (UCL) and Pennsylvania. They claim to have 'trained' (reframed) 20,000 civil servants and run more than 750 projects involving 400 randomised controlled trials in dozens of countries' as another version of mind reframers Common Purpose. BIT works from its office in New York with cities and their agencies, as well as other partners, across the United States and Canada – this is a company part-owned by the British government Cabinet Office. An executive order by President Cult-servant Obama established a US Social and Behavioral Sciences Team in 2015. They all have the same reason for being and that's to brainwash the population directly and by brainwashing those in positions of authority.

'Covid' mind game

Another prime aspect of the UK mind-control network is the 'independent' [joke] Scientific Pandemic Insights Group on Behaviours (SPI-B) which 'provides behavioural science advice aimed at anticipating and helping people adhere to interventions that are recommended by medical or epidemiological experts'. That means manipulating public perception and behaviour to do whatever government tells them to do. It's disgusting and if they really want the public to be 'safe' this lot should all be under lock and key. According to the government website SPI-B consists of

'behavioural scientists, health and social psychologists, anthropologists and historians' and advises the Whitty-Vallance-led Scientific Advisory Group for Emergencies (SAGE) which in turn advises the government on 'the science' (it doesn't) and 'Covid' policy. When politicians say they are being guided by 'the science' this is the rabble in each country they are talking about and that 'science' is dominated by behaviour manipulators to enforce government fascism through public compliance. The Behaviour Insight Team is headed by psychologist David Solomon Halpern, a visiting professor at King's College London, and connects with a national and global web of other civilian and military organisations as the Cult moves towards its goal of fusing them into one fascistic whole in every country through its 'Fusion Doctrine'. The behaviour manipulation network involves, but is not confined to, the Foreign Office; National Security Council; government communications headquarters (GCHQ); MI5; MI6; the Cabinet Office-based Media Monitoring Unit; and the Rapid Response Unit which 'monitors digital trends to spot emerging issues; including misinformation and disinformation; and identifies the best way to respond'.

There is also the 77th Brigade of the UK military which operates like the notorious Israeli military's Unit 8200 in manipulating information and discussion on the Internet by posing as members of the public to promote the narrative and discredit those who challenge it. Here we have the military seeking to manipulate *domestic* public opinion while the Nazis in government are fine with that. Conservative Member of Parliament Tobias Ellwood, an advocate of lockdown and control through 'vaccine passports', is a Lieutenant Colonel reservist in the 77th Brigade which connects with the military operation jHub, the 'innovation centre' for the Ministry of Defence and Strategic Command. jHub has also been involved with the civilian National Health Service (NHS) in 'symptom tracing' the population. The NHS is a key part of this mind control network and produced a document in December, 2020, explaining to staff how to use psychological manipulation with different groups and ages to get them to have the DNA-manipulating 'Covid vaccine'

that's designed to cumulatively rewrite human genetics. The document, called 'Optimising Vaccination Roll Out – Do's and Dont's for all messaging, documents and "communications" in the widest sense', was published by NHS England and the NHS Improvement *Behaviour Change Unit* in partnership with Public Health England and Warwick Business School. I hear the mantra about 'save the NHS' and 'protect the NHS' when we need to scrap the NHS and start again. The current version is far too corrupt, far too anti-human and totally compromised by Cult operatives and their assets. UK government broadcast media censor Ofcom will connect into this web – as will the BBC with its tremendous Ofcom influence – to control what the public see and hear and dictate mass perception. Nuremberg trials must include personnel from all these organisations.

The fear factor

The 'Covid' hoax has led to the creation of the UK Cabinet Office-connected Joint Biosecurity Centre (JBC) which is officially described as providing 'expert advice on pandemics' using its independent [all Cult operations are 'independent'] analytical function to provide real-time analysis about infection outbreaks to identify and respond to outbreaks of Covid-19'. Another role is to advise the government on a response to spikes in infections – 'for example by closing schools or workplaces in local areas where infection levels have risen'. Put another way, promoting the Cult agenda. The Joint Biosecurity Centre is modelled on the Joint Terrorism Analysis Centre which analyses intelligence to set 'terrorism threat levels' and here again you see the fusion of civilian and military operations and intelligence that has led to military intelligence producing documents about 'vaccine hesitancy' and how it can be combated. Domestic civilian matters and opinions should not be the business of the military. The Joint Biosecurity Centre is headed by Tom Hurd, director general of the Office for Security and Counter-Terrorism from the establishment-to-its-fingertips Hurd family. His father is former Foreign Secretary Douglas Hurd. How coincidental that Tom

Hurd went to the elite Eton College and Oxford University with Boris Johnson. Imperial College with its ridiculous computer modeller Neil Ferguson will connect with this gigantic web that will itself interconnect with similar set-ups in other major and not so major countries. Compared with this Cult network the politicians, be they Boris Johnson, Donald Trump or Joe Biden, are bit-part players 'following the science'. The network of psychologists was on the 'Covid' case from the start with the aim of generating maximum fear of the 'virus' to ensure compliance by the population. A government behavioural science group known as SPI-B produced a paper in March, 2020, for discussion by the main government science advisory group known as SAGE. It was headed 'Options for increasing adherence to social distancing measures' and it said the following in a section headed 'Persuasion':

- A substantial number of people still do not feel sufficiently personally threatened; it could be that they are reassured by the low death rate in their demographic group, although levels of concern may be rising. Having a good understanding of the risk has been found to be positively associated with adoption of COVID-19 social distancing measures in Hong Kong.

- The perceived level of personal threat needs to be increased among those who are complacent, using hard-hitting evaluation of options for increasing social distancing emotional messaging. To be effective this must also empower people by making clear the actions they can take to reduce the threat.

- Responsibility to others: There seems to be insufficient understanding of, or feelings of responsibility about, people's role in transmitting the infection to others … Messaging about actions need to be framed positively in terms of protecting oneself and the community, and increase confidence that they will be effective.

- Some people will be more persuaded by appeals to play by the rules, some by duty to the community, and some to personal risk.

> All these different approaches are needed. The messaging also needs to take account of the realities of different people's lives. Messaging needs to take account of the different motivational levers and circumstances of different people.

All this could be achieved the SPI-B psychologists said by *using the media to increase the sense of personal threat* which translates as terrify the shit out of the population, including children, so they all do what we want. That's not happened has it? Those excuses for 'journalists' who wouldn't know journalism if it bit them on the arse (the great majority) have played their crucial part in serving this Cult-government Psyop to enslave their own kids and grandkids. How they live with themselves I have no idea. The psychological war has been underpinned by constant government 'Covid' propaganda in almost every television and radio ad break, plus the Internet and print media, which has pounded out the fear with taxpayers footing the bill for their own programming. The result has been people terrified of a 'virus' that doesn't exist or one with a tiny fatality rate even if you believe it does. People walk down the street and around the shops wearing face-nappies damaging their health and psychology while others report those who refuse to be that naïve to the police who turn up in their own face-nappies. I had a cameraman come to my flat and he was so frightened of 'Covid' he came in wearing a mask and refused to shake my hand in case he caught something. He had – naïveitis – and the thought that he worked in the mainstream media was both depressing and made his behaviour perfectly explainable. The fear which has gripped the minds of so many and frozen them into compliance has been carefully cultivated by these psychologists who are really psychopaths. If lives get destroyed and a lot of young people commit suicide it shows our plan is working. SPI-B then turned to compulsion on the public to comply. 'With adequate preparation, rapid change can be achieved', it said. Some countries had introduced mandatory self-isolation on a wide scale without evidence of major public unrest and a large majority of the UK's population appeared to be supportive of more coercive measures with 64 percent of adults saying they would

support putting London under a lockdown (watch the 'polls' which are designed to make people believe that public opinion is in favour or against whatever the subject in hand).

For 'aggressive protective measures' to be effective, the SPI-B paper said, special attention should be devoted to those population groups that are more at risk. Translated from the Orwellian this means making the rest of population feel guilty for not protecting the 'vulnerable' such as old people which the Cult and its agencies were about to kill on an industrial scale with lockdown, lack of treatment and the Gates 'vaccine'. Psychopath psychologists sold their guilt-trip so comprehensively that Los Angeles County Supervisor Hilda Solis reported that children were apologising (from a distance) to their parents and grandparents for bringing 'Covid' into their homes and getting them sick. '… These apologies are just some of the last words that loved ones will ever hear as they die alone,' she said. Gut-wrenchingly Solis then used this childhood tragedy to tell children to stay at home and 'keep your loved ones alive'. Imagine heaping such potentially life-long guilt on a kid when it has absolutely nothing to do with them. These people are deeply disturbed and the psychologists behind this even more so.

Uncivil war – divide and rule

Professional mind-controllers at SPI-B wanted the media to increase a sense of responsibility to others (do as you're told) and promote 'positive messaging' for those actions while in contrast to invoke 'social disapproval' by the unquestioning, obedient, community of anyone with a mind of their own. Again the compliant Goebbels-like media obliged. This is an old, old, trick employed by tyrannies the world over throughout human history. You get the target population to keep the target population in line – *your* line. SPI-B said this could 'play an important role in preventing anti-social behaviour or discouraging failure to enact pro-social behaviour'. For 'anti-social' in the Orwellian parlance of SPI-B see any behaviour that government doesn't approve. SPI-B recommendations said that 'social disapproval' should be accompanied by clear messaging and

promotion of strong collective identity – hence the government and celebrity mantra of 'we're all in this together'. Sure we are. The mind doctors have such contempt for their targets that they think some clueless comedian, actor or singer telling them to do what the government wants will be enough to win them over. We have had UK comedian Lenny Henry, actor Michael Caine and singer Elton John wheeled out to serve the propagandists by urging people to have the DNA-manipulating 'Covid' non-'vaccine'. The role of Henry and fellow black celebrities in seeking to coax a 'vaccine' reluctant black community into doing the government's will was especially stomach-turning. An emotion-manipulating script and carefully edited video featuring these black 'celebs' was such an insult to the intelligence of black people and where's the self-respect of those involved selling their souls to a fascist government agenda? Henry said he heard black people's 'legitimate worries and concerns', but people must 'trust the facts' when they were doing exactly that by not having the 'vaccine'. They had to include the obligatory reference to Black Lives Matter with the line … 'Don't let coronavirus cost even more black lives – because we matter'. My god, it was pathetic. 'I know the vaccine is safe and what it does.' How? 'I'm a comedian and it says so in my script.'

SPI-B said social disapproval needed to be carefully managed to avoid victimisation, scapegoating and misdirected criticism, but they knew that their 'recommendations' would lead to exactly that and the media were specifically used to stir-up the divide-and-conquer hostility. Those who conform like good little baa, baas, are praised while those who have seen through the tidal wave of lies are 'Covidiots'. The awake have been abused by the fast asleep for not conforming to fascism and impositions that the awake know are designed to endanger their health, dehumanise them, and tear asunder the very fabric of human society. We have had the curtain-twitchers and morons reporting neighbours and others to the face-nappied police for breaking 'Covid rules' with fascist police delighting in posting links and phone numbers where this could be done. The Cult cannot impose its will without a compliant police

and military or a compliant population willing to play their part in enslaving themselves and their kids. The words of a pastor in Nazi Germany are so appropriate today:

First they came for the socialists and I did not speak out because I was not a socialist.

Then they came for the trade unionists and I did not speak out because I was not a trade unionist.

Then they came for the Jews and I did not speak out because I was not a Jew.

Then they came for me and there was no one left to speak for me.

Those who don't learn from history are destined to repeat it and so many are.

'Covid' rules: Rewiring the mind

With the background laid out to this gigantic national and global web of psychological manipulation we can put 'Covid' rules into a clear and sinister perspective. Forget the claims about protecting health. 'Covid' rules are about dismantling the human mind, breaking the human spirit, destroying self-respect, and then putting Humpty Dumpty together again as a servile, submissive slave. Social isolation through lockdown and distancing have devastating effects on the human psyche as the psychological psychopaths well know and that's the real reason for them. Humans need contact with each other, discourse, closeness and touch, or they eventually, and literarily, go crazy. Masks, which I will address at some length, fundamentally add to the effects of isolation and the Cult agenda to dehumanise and de-individualise the population. To do this while knowing – in fact *seeking* – this outcome is the very epitome of evil and psychologists involved in this *are* the epitome of evil. They must like all the rest of the Cult demons and their assets stand trial for crimes against humanity on a scale that defies the imagination. Psychopaths in uniform use isolation to break enemy troops and agents and make them subservient and submissive to tell what they know. The technique is rightly considered a form of torture and

torture is most certainly what has been imposed on the human population.

Clinically-insane American psychologist Harry Harlow became famous for his isolation experiments in the 1950s in which he separated baby monkeys from their mothers and imprisoned them for months on end in a metal container or 'pit of despair'. They soon began to show mental distress and depression as any idiot could have predicted. Harlow put other monkeys in steel chambers for three, six or twelve months while denying them any contact with animals or humans. He said that the effects of total social isolation for six months were 'so devastating and debilitating that we had assumed initially that twelve months of isolation would not produce any additional decrement'; but twelve months of isolation 'almost obliterated the animals socially'. This is what the Cult and its psychopaths are doing to you and your children. Even monkeys in partial isolation in which they were not allowed to form relationships with other monkeys became 'aggressive and hostile, not only to others, but also towards their own bodies'. We have seen this in the young as a consequence of lockdown. UK government psychopaths launched a public relations campaign telling people not to hug each other even after they received the 'Covid-19 vaccine' which we were told with more lies would allow a return to 'normal life'. A government source told *The Telegraph*: 'It will be along the lines that it is great that you have been vaccinated, but if you are going to visit your family and hug your grandchildren there is a chance you are going to infect people you love.' The source was apparently speaking from a secure psychiatric facility. Janet Lord, director of Birmingham University's Institute of Inflammation and Ageing, said that parents and grandparents should avoid hugging their children. Well, how can I put it, Ms Lord? Fuck off. Yep, that'll do.

Destroying the kids – where are the parents?

Observe what has happened to people enslaved and isolated by lockdown as suicide and self-harm has soared worldwide,

particularly among the young denied the freedom to associate with their friends. A study of 49,000 people in English-speaking countries concluded that almost half of young adults are at clinical risk of mental health disorders. A national survey in America of 1,000 currently enrolled high school and college students found that 5 percent reported attempting suicide during the pandemic. Data from the US CDC's National Syndromic Surveillance Program from January 1st to October 17th, 2020, revealed a *31 percent* increase in mental health issues among adolescents aged 12 to 17 compared with 2019. The CDC reported that America in general suffered the biggest drop in life expectancy since World War Two as it fell by a year in the first half of 2020 as a result of 'deaths of despair' – overdoses and suicides. Deaths of despair have leapt by more than 20 percent during lockdown and include the highest number of fatal overdoses ever recorded in a single year – 81,000. Internet addiction is another consequence of being isolated at home which lowers interest in physical activities as kids fall into inertia and what's the point? Children and young people are losing hope and giving up on life, sometimes literally. A 14-year-old boy killed himself in Maryland because he had 'given up' when his school district didn't reopen; an 11-year-old boy shot himself during a zoom class; a teenager in Maine succumbed to the isolation of the 'pandemic' when he ended his life after experiencing a disrupted senior year at school. Children as young as nine have taken their life and all these stories can be repeated around the world. Careers are being destroyed before they start and that includes those in sport in which promising youngsters have not been able to take part. The plan of the psycho-psychologists is working all right. Researchers at Cambridge University found that lockdowns cause significant harm to children's mental health. Their study was published in the *Archives of Disease in Childhood*, and followed 168 children aged between 7 and 11. The researchers concluded:

> During the UK lockdown, children's depression symptoms have increased substantially, relative to before lockdown. The scale of this effect has direct relevance for the continuation of different elements of lockdown policy, such as complete or partial school closures …

… Specifically, we observed a statistically significant increase in ratings of depression, with a medium-to-large effect size. Our findings emphasise the need to incorporate the potential impact of lockdown on child mental health in planning the ongoing response to the global pandemic and the recovery from it.

Not a chance when the Cult's psycho-psychologists were getting exactly what they wanted. The UK's Royal College of Paediatrics and Child Health has urged parents to look for signs of eating disorders in children and young people after a three to four fold increase. Specialists say the 'pandemic' is a major reason behind the rise. You don't say. The College said isolation from friends during school closures, exam cancellations, loss of extra-curricular activities like sport, and an increased use of social media were all contributory factors along with fears about the virus (psycho-psychologists again), family finances, and students being forced to quarantine. Doctors said young people were becoming severely ill by the time they were seen with 'Covid' regulations reducing face-to-face consultations. Nor is it only the young that have been devastated by the psychopaths. Like all bullies and cowards the Cult is targeting the young, elderly, weak and infirm. A typical story was told by a British lady called Lynn Parker who was not allowed to visit her husband in 2020 for the last ten and half months of his life 'when he needed me most' between March 20th and when he died on December 19th. This vacates the criminal and enters the territory of evil. The emotional impact on the immune system alone is immense as are the number of people of all ages worldwide who have died as a result of Cult-demanded, Gates-demanded, lockdowns.

Isolation is torture

The experience of imposing solitary confinement on millions of prisoners around the world has shown how a large percentage become 'actively psychotic and/or acutely suicidal'. Social isolation has been found to trigger 'a specific psychiatric syndrome, characterized by hallucinations; panic attacks; overt paranoia; diminished impulse control; hypersensitivity to external stimuli; and difficulties with thinking, concentration and memory'. Juan Mendez,

a United Nations rapporteur (investigator), said that isolation is a form of torture. Research has shown that even after isolation prisoners find it far more difficult to make social connections and I remember chatting to a shop assistant after one lockdown who told me that when her young son met another child again he had no idea how to act or what to do. Hannah Flanagan, Director of Emergency Services at Journey Mental Health Center in Dane County, Wisconsin, said: 'The specificity about Covid social distancing and isolation that we've come across as contributing factors to the suicides are really new to us this year.' But they are not new to those that devised them. They are getting the effect they want as the population is psychologically dismantled to be rebuilt in a totally different way. Children and the young are particularly targeted. They will be the adults when the full-on fascist AI-controlled technocracy is planned to be imposed and they are being prepared to meekly submit. At the same time older people who still have a memory of what life was like before – and how fascist the new normal really is – are being deleted. You are going to see efforts to turn the young against the old to support this geriatric genocide. Hannah Flanagan said the big increase in suicide in her county proved that social isolation is not only harmful, but deadly. Studies have shown that isolation from others is one of the main risk factors in suicide and even more so with women. Warnings that lockdown could create a 'perfect storm' for suicide were ignored. After all this was one of the *reasons* for lockdown. Suicide, however, is only the most extreme of isolation consequences. There are many others. Dr Dhruv Khullar, assistant professor of healthcare policy at Weill Cornell Medical College, said in a *New York Times* article in 2016 long before the fake 'pandemic':

> A wave of new research suggests social separation is bad for us. Individuals with less social connection have disrupted sleep patterns, altered immune systems, more inflammation and higher levels of stress hormones. One recent study found that isolation increases the risk of heart disease by 29 percent and stroke by 32 percent. Another analysis that pooled data from 70 studies and 3.4 million people found that socially isolated individuals had a 30 percent higher risk of dying in the next seven years, and that this effect was largest in middle age.

Loneliness can accelerate cognitive decline in older adults, and isolated individuals are twice as likely to die prematurely as those with more robust social interactions. These effects start early: Socially isolated children have significantly poorer health 20 years later, even after controlling for other factors. All told, loneliness is as important a risk factor for early death as obesity and smoking.

There you have proof from that one article alone four years before 2020 that those who have enforced lockdown, social distancing and isolation knew what the effect would be and that is even more so with professional psychologists that have been driving the policy across the globe. We can go back even further to the years 2000 and 2003 and the start of a major study on the effects of isolation on health by Dr Janine Gronewold and Professor Dirk M. Hermann at the University Hospital in Essen, Germany, who analysed data on 4,316 people with an average age of 59 who were recruited for the long-term research project. They found that socially isolated people are more than 40 percent more likely to have a heart attack, stroke, or other major cardiovascular event and nearly 50 percent more likely to die from any cause. Given the financial Armageddon unleashed by lockdown we should note that the study found a relationship between increased cardiovascular risk and lack of financial support. After excluding other factors social isolation was still connected to a 44 percent increased risk of cardiovascular problems and a 47 percent increased risk of death by any cause. Lack of financial support was associated with a 30 percent increase in the risk of cardiovascular health events. Dr Gronewold said it had been known for some time that feeling lonely or lacking contact with close friends and family can have an impact on physical health and the study had shown that having strong social relationships is of high importance for heart health. Gronewold said they didn't understand yet why people who are socially isolated have such poor health outcomes, but this was obviously a worrying finding, particularly during these times of prolonged social distancing. Well, it can be explained on many levels. You only have to identify the point in the body where people feel loneliness and missing people they are parted from – it's in the centre of the chest where they feel the ache of loneliness and the ache of missing people. 'My heart aches for

you' … 'My heart aches for some company.' I will explain this more in the chapter Escaping Wetiko, but when you realise that the body is the mind – they are expressions of each other – the reason why state of the mind dictates state of the body becomes clear.

American psychologist Ranjit Powar was highlighting the effects of lockdown isolation as early as April, 2020. She said humans have evolved to be social creatures and are wired to live in interactive groups. Being isolated from family, friends and colleagues could be unbalancing and traumatic for most people and could result in short or even long-term psychological and physical health problems. An increase in levels of anxiety, aggression, depression, forgetfulness and hallucinations were possible psychological effects of isolation. 'Mental conditions may be precipitated for those with underlying pre-existing susceptibilities and show up in many others without any pre-condition.' Powar said personal relationships helped us cope with stress and if we lost this outlet for letting off steam the result can be a big emotional void which, for an average person, was difficult to deal with. 'Just a few days of isolation can cause increased levels of anxiety and depression' – so what the hell has been the effect on the global population of *18 months* of this at the time of writing? Powar said: 'Add to it the looming threat of a dreadful disease being repeatedly hammered in through the media and you have a recipe for many shades of mental and physical distress.' For those with a house and a garden it is easy to forget that billions have had to endure lockdown isolation in tiny overcrowded flats and apartments with nowhere to go outside. The psychological and physical consequences of this are unimaginable and with lunatic and abusive partners and parents the consequences have led to tremendous increases in domestic and child abuse and alcoholism as people seek to shut out the horror. Ranjit Powar said:

Staying in a confined space with family is not all a rosy picture for everyone. It can be extremely oppressive and claustrophobic for large low-income families huddled together in small single-room houses. Children here are not lucky enough to have many board/electronic games or books to keep them occupied.

Add to it the deep insecurity of running out of funds for food and basic necessities. On the other hand, there are people with dysfunctional family dynamics, such as domineering, abusive or alcoholic partners, siblings or parents which makes staying home a period of trial. Incidence of suicide and physical abuse against women has shown a worldwide increase. Heightened anxiety and depression also affect a person's immune system, making them more susceptible to illness.

To think that Powar's article was published on April 11th, 2020.

Six-feet fantasy

Social (unsocial) distancing demanded that people stay six feet or two metres apart. UK government advisor Robert Dingwall from the New and Emerging Respiratory Virus Threats Advisory Group said in a radio interview that the two-metre rule was 'conjured up out of nowhere' and was not based on science. No, it was not based on *medical* science, but it didn't come out of nowhere. The distance related to *psychological* science. Six feet/two metres was adopted in many countries and we were told by people like the criminal Anthony Fauci and his ilk that it was founded on science. Many schools could not reopen because they did not have the space for six-feet distancing. Then in March, 2021, after a year of six-feet 'science', a study published in the *Journal of Infectious Diseases* involving more than 500,000 students and almost 100,000 staff over 16 weeks revealed no significant difference in 'Covid' cases between six feet and three feet and Fauci changed his tune. Now three feet was okay. There is no difference between six feet and three *inches* when there is no 'virus' and they got away with six feet for psychological reasons for as long as they could. I hear journalists and others talk about 'unintended consequences' of lockdown. They are not *unintended* at all; they have been coldly-calculated for a specific outcome of human control and that's why super-psychopaths like Gates have called for them so vehemently. Super-psychopath psychologists have demanded them and psychopathic or clueless, spineless, politicians have gone along with them by 'following the science'. But it's not science at all. 'Science' is not what is; it's only what people can be manipulated to believe it is. The whole 'Covid' catastrophe is

founded on mind control. Three word or three statement mantras issued by the UK government are a well-known mind control technique and so we've had 'Stay home/protect the NHS/save lives', 'Stay alert/control the virus/save lives' and 'hands/face/space'. One of the most vocal proponents of extreme 'Covid' rules in the UK has been Professor Susan Michie, a member of the British Communist Party, who is not a medical professional. Michie is the director of the Centre for Behaviour Change at University College London. She is a *behavioural psychologist* and another filthy rich 'Marxist' who praised China's draconian lockdown. She was known by fellow students at Oxford University as 'Stalin's nanny' for her extreme Marxism. Michie is an influential member of the UK government's Scientific Advisory Group for Emergencies (SAGE) and behavioural manipulation groups which have dominated 'Covid' policy. She is a consultant adviser to the World Health Organization on 'Covid-19' and behaviour. Why the hell are lockdowns anything to do with her when they are claimed to be about health? Why does a behavioural psychologist from a group charged with changing the behaviour of the public want lockdown, human isolation and mandatory masks? Does that question really need an answer? Michie *absolutely* has to explain herself before a Nuremberg court when humanity takes back its world again and even more so when you see the consequences of masks that she demands are compulsory. This is a Michie classic:

> The benefits of getting primary school children to wear masks is that regardless of what little degree of transmission is occurring in those age groups it could help normalise the practice. Young children wearing masks may be more likely to get their families to accept masks.

Those words alone should carry a prison sentence when you ponder on the callous disregard for children involved and what a statement it makes about the mind and motivations of Susan Michie. What a lovely lady and what she said there encapsulates the mentality of the psychopaths behind the 'Covid' horror. Let us compare what Michie said with a countrywide study in Germany published at researchsquare.com involving 25,000 school children and 17,854 health complaints submitted by parents. Researchers

found that masks are harming children physically, psychologically, and behaviourally with 24 health issues associated with mask wearing. They include: shortness of breath (29.7%); dizziness (26.4%); increased headaches (53%); difficulty concentrating (50%); drowsiness or fatigue (37%); and malaise (42%). Nearly a third of children experienced more sleep issues than before and a quarter developed new fears. Researchers found health issues and other impairments in 68 percent of masked children covering their faces for an average of 4.5 hours a day. Hundreds of those taking part experienced accelerated respiration, tightness in the chest, weakness, and short-term impairment of consciousness. A reminder of what Michie said again:

The benefits of getting primary school children to wear masks is that regardless of what little degree of transmission is occurring in those age groups it could help normalise the practice. Young children wearing masks may be more likely to get their families to accept masks.

Psychopaths in government and psychology now have children and young people – plus all the adults – wearing masks for hours on end while clueless teachers impose the will of the psychopaths on the young they should be protecting. What the hell are parents doing?

Cult lab rats

We have some schools already imposing on students microchipped buzzers that activate when they get 'too close' to their pals in the way they do with lab rats. How apt. To the Cult and its brain-dead servants our children *are* lab rats being conditioned to be unquestioning, dehumanised slaves for the rest of their lives. Children and young people are being weaned and frightened away from the most natural human instincts including closeness and touch. I have tracked in the books over the years how schools were banning pupils from greeting each other with a hug and the whole Cult-induced Me Too movement has terrified men and boys from a relaxed and natural interaction with female friends and work colleagues to the point where many men try never to be in a room

alone with a woman that's not their partner. Airhead celebrities have as always played their virtue-signalling part in making this happen with their gross exaggeration. For every monster like Harvey Weinstein there are at least tens of thousands of men that don't treat women like that; but everyone must be branded the same and policy changed for them as well as the monster. I am going to be using the word 'dehumanise' many times in this chapter because that is what the Cult is seeking to do and it goes very deep as we shall see. Don't let them kid you that social distancing is planned to end one day. That's not the idea. We are seeing more governments and companies funding and producing wearable gadgets to keep people apart and they would not be doing that if this was meant to be short-term. A tech start-up company backed by GCHQ, the British Intelligence and military surveillance headquarters, has created a social distancing wrist sensor that alerts people when they get too close to others. The CIA has also supported tech companies developing similar devices. The wearable sensor was developed by Tended, one of a number of start-up companies supported by GCHQ (see the CIA and DARPA). The device can be worn on the wrist or as a tag on the waistband and will vibrate whenever someone wearing the device breaches social distancing and gets anywhere near natural human contact. The company had a lucky break in that it was developing a distancing sensor when the 'Covid' hoax arrived which immediately provided a potentially enormous market. How fortunate. The government in big-time Cult-controlled Ontario in Canada is investing $2.5 million in wearable contact tracing technology that 'will alert users if they may have been exposed to the Covid-19 in the workplace and will beep or vibrate if they are within six feet of another person'. Facedrive Inc., the technology company behind this, was founded in 2016 with funding from the Ontario Together Fund and obviously they, too, had a prophet on the board of directors. The human surveillance and control technology is called TraceSCAN and would be worn by the human cyborgs in places such as airports, workplaces, construction sites, care homes and … *schools*.

I emphasise schools with children and young people the prime targets. You know what is planned for society as a whole if you keep your eyes on the schools. They have always been places where the state program the next generation of slaves to be its compliant worker-ants – or Woker-ants these days; but in the mist of the 'Covid' madness they have been transformed into mind laboratories on a scale never seen before. Teachers and head teachers are just as programmed as the kids – often more so. Children are kept apart from human interaction by walk lanes, classroom distancing, staggered meal times, masks, and the rolling-out of buzzer systems. Schools are now physically laid out as a laboratory maze for lab-rats. Lunatics at a school in Anchorage, Alaska, who should be prosecuted for child abuse, took away desks and forced children to kneel (know your place) on a mat for five hours a day while wearing a mask and using their chairs as a desk. How this was supposed to impact on a 'virus' only these clinically insane people can tell you and even then it would be clap-trap. The school banned recess (interaction), art classes (creativity), and physical exercise (getting body and mind moving out of inertia). Everyone behind this outrage should be in jail or better still a mental institution. The behavioural manipulators are all for this dystopian approach to schools. Professor Susan Michie, the mind-doctor and British Communist Party member, said it was wrong to say that schools were safe. They had to be made so by 'distancing', masks and ventilation (sitting all day in the cold). I must ask this lady round for dinner on a night I know I am going to be out and not back for weeks. She probably wouldn't be able to make it, anyway, with all the visits to her own psychologist she must have block-booked.

Masking identity

I know how shocking it must be for you that a behaviour manipulator like Michie wants everyone to wear masks which have long been a feature of mind-control programs like the infamous MKUltra in the United States, but, there we are. We live and learn. I spent many years from 1996 to right across the millennium

researching mind control in detail on both sides of the Atlantic and elsewhere. I met a large number of mind-control survivors and many had been held captive in body and mind by MKUltra. MK stands for mind-control, but employs the German spelling in deference to the Nazis spirited out of Germany at the end of World War Two by Operation Paperclip in which the US authorities, with help from the Vatican, transported Nazi mind-controllers and engineers to America to continue their work. Many of them were behind the creation of NASA and they included Nazi scientist and SS officer Wernher von Braun who swapped designing V-2 rockets to bombard London with designing the Saturn V rockets that powered the NASA moon programme's Apollo craft. I think I may have mentioned that the Cult has no borders. Among Paperclip escapees was Josef Mengele, the Angel of Death in the Nazi concentration camps where he conducted mind and genetic experiments on children often using twins to provide a control twin to measure the impact of his 'work' on the other. If you want to observe the Cult mentality in all its extremes of evil then look into the life of Mengele. I have met many people who suffered mercilessly under Mengele in the United States where he operated under the name Dr Greene and became a stalwart of MKUltra programming and torture. Among his locations was the underground facility in the Mojave Desert in California called the China Lake Naval Weapons Station which is almost entirely below the surface. My books *The Biggest Secret*, *Children of the Matrix* and *The Perception Deception* have the detailed background to MKUltra.

The best-known MKUltra survivor is American Cathy O'Brien. I first met her and her late partner Mark Phillips at a conference in Colorado in 1996. Mark helped her escape and deprogram from decades of captivity in an offshoot of MKUltra known as Project Monarch in which 'sex slaves' were provided for the rich and famous including Father George Bush, Dick Cheney and the Clintons. Read Cathy and Mark's book *Trance-Formation of America* and if you are new to this you will be shocked to the core. I read it in 1996 shortly before, with the usual synchronicity of my life, I found

myself given a book table at the conference right next to hers. MKUltra never ended despite being very publicly exposed (only a small part of it) in the 1970s and continues in other guises. I am still in touch with Cathy. She contacted me during 2020 after masks became compulsory in many countries to tell me how they were used as part of MKUltra programming. I had been observing 'Covid regulations' and the relationship between authority and public for months. I saw techniques that I knew were employed on individuals in MKUltra being used on the global population. I had read many books and manuals on mind control including one called *Silent Weapons for Quiet Wars* which came to light in the 1980s and was a guide on how to perceptually program on a mass scale. 'Silent Weapons' refers to mind-control. I remembered a line from the manual as governments, medical authorities and law enforcement agencies have so obviously talked to – or rather at – the adult population since the 'Covid' hoax began as if they are children. The document said:

> If a person is spoken to by a T.V. advertiser as if he were a twelve-year-old, then, due to suggestibility, he will, with a certain probability, respond or react to that suggestion with the uncritical response of a twelve-year-old and will reach in to his economic reservoir and deliver its energy to buy that product on impulse when he passes it in the store.

That's why authority has spoken to adults like children since all this began.

Why did Michael Jackson wear masks?

Every aspect of the 'Covid' narrative has mind-control as its central theme. Cathy O'Brien wrote an article for davidicke.com about the connection between masks and mind control. Her daughter Kelly who I first met in the 1990s was born while Cathy was still held captive in MKUltra. Kelly was forced to wear a mask as part of her programming from the age of *two* to dehumanise her, target her sense of individuality and reduce the amount of oxygen her brain and body received. *Bingo*. This is the real reason for compulsory

masks, why they have been enforced en masse, and why they seek to increase the number they demand you wear. First one, then two, with one disgraceful alleged 'doctor' recommending four which is nothing less than a death sentence. Where and how often they must be worn is being expanded for the purpose of mass mind control and damaging respiratory health which they can call 'Covid-19'. Canada's government headed by the man-child Justin Trudeau, says it's fine for children of two and older to wear masks. An insane 'study' in Italy involving just 47 children concluded there was no problem for babies as young as *four months* wearing them. Even after people were 'vaccinated' they were still told to wear masks by the criminal that is Anthony Fauci. Cathy wrote that mandating masks is allowing the authorities literally to control the air we breathe which is what was done in MKUltra. You might recall how the singer Michael Jackson wore masks and there is a reason for that. He was subjected to MKUltra mind control through Project Monarch and his psyche was scrambled by these simpletons. Cathy wrote:

In MKUltra Project Monarch mind control, Michael Jackson had to wear a mask to silence his voice so he could not reach out for help. Remember how he developed that whisper voice when he wasn't singing? Masks control the mind from the outside in, like the redefining of words is doing. By controlling what we can and cannot say for fear of being labeled racist or beaten, for example, it ultimately controls thought that drives our words and ultimately actions (or lack thereof).

Likewise, a mask muffles our speech so that we are not heard, which controls voice … words … mind. This is Mind Control. Masks are an obvious mind control device, and I am disturbed so many people are complying on a global scale. Masks depersonalize while making a person feel as though they have no voice. It is a barrier to others. People who would never choose to comply but are forced to wear a mask in order to keep their job, and ultimately their family fed, are compromised. They often feel shame and are subdued. People have stopped talking with each other while media controls the narrative.

The 'no voice' theme has often become literal with train passengers told not to speak to each other in case they pass on the 'virus', singing banned for the same reason and bonkers California officials telling people riding roller coasters that they cannot shout and scream. Cathy said she heard every day from healed MKUltra survivors who cannot wear a mask without flashing back on ways

their breathing was controlled – 'from ball gags and penises to water boarding'. She said that through the years when she saw images of people in China wearing masks 'due to pollution' that it was really to control their oxygen levels. 'I knew it was as much of a population control mechanism of depersonalisation as are burkas', she said. Masks are another Chinese communist/fascist method of control that has been swept across the West as the West becomes China at lightning speed since we entered 2020.

Mask-19

There are other reasons for mandatory masks and these include destroying respiratory health to call it 'Covid-19' and stunting brain development of children and the young. Dr Margarite Griesz-Brisson MD, PhD, is a Consultant Neurologist and Neurophysiologist and the Founder and Medical Director of the London Neurology and Pain Clinic. Her CV goes down the street and round the corner. She is clearly someone who cares about people and won't parrot the propaganda. Griesz-Brisson has a PhD in pharmacology, with special interest in neurotoxicology, environmental medicine, neuroregeneration and neuroplasticity (the way the brain can change in the light of information received). She went public in October, 2020, with a passionate warning about the effects of mask-wearing laws:

> The reinhalation of our exhaled air will without a doubt create oxygen deficiency and a flooding of carbon dioxide. We know that the human brain is very sensitive to oxygen deprivation. There are nerve cells for example in the hippocampus that can't be longer than 3 minutes without oxygen – they cannot survive. The acute warning symptoms are headaches, drowsiness, dizziness, issues in concentration, slowing down of reaction time – reactions of the cognitive system.

Oh, I know, let's tell bus, truck and taxi drivers to wear them and people working machinery. How about pilots, doctors and police? Griesz-Brisson makes the important point that while the symptoms she mentions may fade as the body readjusts this does not alter the fact that people continue to operate in oxygen deficit with long list of

potential consequences. She said it was well known that neurodegenerative diseases take years or decades to develop. 'If today you forget your phone number, the breakdown in your brain would have already started 20 or 30 years ago.' She said degenerative processes in your brain are getting amplified as your oxygen deprivation continues through wearing a mask. Nerve cells in the brain are unable to divide themselves normally in these circumstances and lost nerve cells will no longer be regenerated. 'What is gone is gone.' Now consider that people like shop workers and *schoolchildren* are wearing masks for hours every day. What in the name of sanity is going to be happening to them? 'I do not wear a mask, I need my brain to think', Griesz-Brisson said, 'I want to have a clear head when I deal with my patients and not be in a carbon dioxide-induced anaesthesia'. If you are told to wear a mask anywhere ask the organisation, police, store, whatever, for their risk assessment on the dangers and negative effects on mind and body of enforcing mask-wearing. They won't have one because it has never been done not even by government. All of them must be subject to class-action lawsuits as the consequences come to light. They don't do mask risk assessments for an obvious reason. They know what the conclusions would be and independent scientific studies that *have* been done tell a horror story of consequences.

'Masks are criminal'

Dr Griesz-Brisson said that for children and adolescents, masks are an absolute no-no. They had an extremely active and adaptive immune system and their brain was incredibly active with so much to learn. 'The child's brain, or the youth's brain, is thirsting for oxygen.' The more metabolically active an organ was, the more oxygen it required; and in children and adolescents every organ was metabolically active. Griesz-Brisson said that to deprive a child's or adolescent's brain of oxygen, or to restrict it in any way, was not only dangerous to their health, it was absolutely criminal. 'Oxygen deficiency inhibits the development of the brain, and the damage that has taken place as a result CANNOT be reversed.' Mind

manipulators of MKUltra put masks on two-year-olds they wanted to neurologically rewire and you can see why. Griesz-Brisson said a child needs the brain to learn and the brain needs oxygen to function. 'We don't need a clinical study for that. This is simple, indisputable physiology.' Consciously and purposely induced oxygen deficiency was an absolutely deliberate health hazard, and an absolute medical contraindication which means that 'this drug, this therapy, this method or measure should not be used, and is not allowed to be used'. To coerce an entire population to use an absolute medical contraindication by force, she said, there had to be definite and serious reasons and the reasons must be presented to competent interdisciplinary and independent bodies to be verified and authorised. She had this warning of the consequences that were coming if mask wearing continued:

When, in ten years, dementia is going to increase exponentially, and the younger generations couldn't reach their god-given potential, it won't help to say 'we didn't need the masks'. I know how damaging oxygen deprivation is for the brain, cardiologists know how damaging it is for the heart, pulmonologists know how damaging it is for the lungs. Oxygen deprivation damages every single organ. Where are our health departments, our health insurance, our medical associations? It would have been their duty to be vehemently against the lockdown and to stop it and stop it from the very beginning.

Why do the medical boards issue punishments to doctors who give people exemptions? Does the person or the doctor seriously have to prove that oxygen deprivation harms people? What kind of medicine are our doctors and medical associations representing? Who is responsible for this crime? The ones who want to enforce it? The ones who let it happen and play along, or the ones who don't prevent it?

All of the organisations and people she mentions there either answer directly to the Cult or do whatever hierarchical levels above them tell them to do. The outcome of both is the same. 'It's not about masks, it's not about viruses, it's certainly not about your health', Griesz-Brisson said. 'It is about much, much more. I am not participating. I am not afraid.' They were taking our air to breathe and there was no unfounded medical exemption from face masks. Oxygen deprivation was dangerous for every single brain. It had to be the free decision of every human being whether they want to

wear a mask that was absolutely ineffective to protect themselves from a virus. She ended by rightly identifying where the responsibility lies for all this:

> The imperative of the hour is personal responsibility. We are responsible for what we think, not the media. We are responsible for what we do, not our superiors. We are responsible for our health, not the World Health Organization. And we are responsible for what happens in our country, not the government.

Halle-bloody-lujah.

But surgeons wear masks, right?

Independent studies of mask-wearing have produced a long list of reports detailing mental, emotional and physical dangers. What a definition of insanity to see police officers imposing mask-wearing on the public which will cumulatively damage their health while the police themselves wear masks that will cumulatively damage *their* health. It's utter madness and both public and police do this because 'the government says so' – yes a government of brain-donor idiots like UK Health Secretary Matt Hancock reading the 'follow the science' scripts of psychopathic, lunatic psychologists. The response you get from Stockholm syndrome sufferers defending the very authorities that are destroying them and their families is that 'surgeons wear masks'. This is considered the game, set and match that they must work and don't cause oxygen deficit. Well, actually, scientific studies have shown that they *do* and oxygen levels are monitored in operating theatres to compensate. Surgeons wear masks to stop spittle and such like dropping into open wounds – not to stop 'viral particles' which are so miniscule they can only be seen through an electron microscope. Holes in the masks are significantly bigger than 'viral particles' and if you sneeze or cough they will breach the mask. I watched an incredibly disingenuous 'experiment' that claimed to prove that masks work in catching 'virus' material from the mouth and nose. They did this with a slow motion camera and the mask did block big stuff which stayed inside the mask and

against the face to be breathed in or cause infections on the face as we have seen with many children. 'Viral particles', however, would never have been picked up by the camera as they came through the mask when they are far too small to be seen. The 'experiment' was therefore disingenuous *and* useless.

Studies have concluded that wearing masks in operating theatres (and thus elsewhere) make no difference to preventing infection while the opposite is true with toxic shite building up in the mask and this had led to an explosion in tooth decay and gum disease dubbed by dentists 'mask mouth'. You might have seen the Internet video of a furious American doctor urging people to take off their masks after a four-year-old patient had been rushed to hospital the night before and nearly died with a lung infection that doctors sourced to mask wearing. A study in the journal *Cancer Discovery* found that inhalation of harmful microbes can contribute to advanced stage lung cancer in adults and long-term use of masks can help breed dangerous pathogens. Microbiologists have said frequent mask wearing creates a moist environment in which microbes can grow and proliferate before entering the lungs. The Canadian Agency for Drugs and Technologies in Health, or CADTH, a Canadian national organisation that provides research and analysis to healthcare decision-makers, said this as long ago as 2013 in a report entitled 'Use of Surgical Masks in the Operating Room: A Review of the Clinical Effectiveness and Guidelines'. It said:

- No evidence was found to support the use of surgical face masks to reduce the frequency of surgical site infections
- No evidence was found on the effectiveness of wearing surgical face masks to protect staff from infectious material in the operating room.
- Guidelines recommend the use of surgical face masks by staff in the operating room to protect both operating room staff and patients (despite the lack of evidence).

We were told that the world could go back to 'normal' with the arrival of the 'vaccines'. When they came, fraudulent as they are, the story changed as I knew that it would. We are in the midst of transforming 'normal', not going back to it. Mary Ramsay, head of immunisation at Public Health England, echoed the words of US criminal Anthony Fauci who said masks and other regulations must stay no matter if people are vaccinated. The Fauci idiot continued to wear two masks – different colours so both could be clearly seen – after he *claimed* to have been vaccinated. Senator Rand Paul told Fauci in one exchange that his double-masks were 'theatre' and he was right. It's all theatre. Mary Ramsay back-tracked on the vaccine-return-to-normal theme when she said the public may need to wear masks and social-distance for years despite the jabs. 'People have got used to those lower-level restrictions now, and [they] can live with them', she said telling us what the idea has been all along. 'The vaccine does not give you a pass, even if you have had it, you must continue to follow all the guidelines' said a Public Health England statement which reneged on what we had been told before and made having the 'vaccine' irrelevant to 'normality' even by the official story. Spain's fascist government trumped everyone by passing a law mandating the wearing of masks on the beach and even when swimming in the sea. The move would have devastated what's left of the Spanish tourist industry, posed potential breathing dangers to swimmers and had Northern European sunbathers walking around with their forehead brown and the rest of their face white as a sheet. The ruling was so crazy that it had to be retracted after pressure from public and tourist industry, but it confirmed where the Cult wants to go with masks and how clinically insane authority has become. The determination to make masks permanent and hide the serious dangers to body and mind can be seen in the censorship of scientist Professor Denis Rancourt by Bill Gates-funded academic publishing website ResearchGate over his papers exposing the dangers and uselessness of masks. Rancourt said:

ResearchGate today has permanently locked my account, which I have had since 2015. Their reasons graphically show the nature of their attack against democracy, and their corruption of

science … By their obscene non-logic, a scientific review of science articles reporting on harms caused by face masks has a 'potential to cause harm'. No criticism of the psychological device (face masks) is tolerated, if the said criticism shows potential to influence public policy.

This is what happens in a fascist world.

Where are the 'greens' (again)?

Other dangers of wearing masks especially regularly relate to the inhalation of minute plastic fibres into the lungs and the deluge of discarded masks in the environment and oceans. Estimates predicted that more than 1.5 billion disposable masks will end up in the world's oceans every year polluting the water with tons of plastic and endangering marine wildlife. Studies project that humans are using 129 billion face masks each month worldwide – about three million a minute. Most are disposable and made from plastic, non-biodegradable microfibers that break down into smaller plastic particles that become widespread in ecosystems. They are littering cities, clogging sewage channels and turning up in bodies of water. I have written in other books about the immense amounts of microplastics from endless sources now being absorbed into the body. Rolf Halden, director of the Arizona State University (ASU) Biodesign Center for Environmental Health Engineering, was the senior researcher in a 2020 study that analysed 47 human tissue samples and found microplastics in all of them. 'We have detected these chemicals of plastics in every single organ that we have investigated', he said. I wrote in *The Answer* about the world being deluged with microplastics. A study by the Worldwide Fund for Nature (WWF) found that people are consuming on average every week some 2,000 tiny pieces of plastic mostly through water and also through marine life and the air. Every year humans are ingesting enough microplastics to fill a heaped dinner plate and in a life-time of 79 years it is enough to fill two large waste bins. Marco Lambertini, WWF International director general said: 'Not only are plastics polluting our oceans and waterways and killing marine life – it's in all of us and we can't escape consuming plastics,' American

geologists found tiny plastic fibres, beads and shards in rainwater samples collected from the remote slopes of the Rocky Mountain National Park near Denver, Colorado. Their report was headed: 'It is raining plastic.' Rachel Adams, senior lecturer in Biomedical Science at Cardiff Metropolitan University, said that among health consequences are internal inflammation and immune responses to a 'foreign body'. She further pointed out that microplastics become carriers of toxins including mercury, pesticides and dioxins (a known cause of cancer and reproductive and developmental problems). These toxins accumulate in the fatty tissues once they enter the body through microplastics. Now this is being compounded massively by people putting plastic on their face and throwing it away.

Workers exposed to polypropylene plastic fibres known as 'flock' have developed 'flock worker's lung' from inhaling small pieces of the flock fibres which can damage lung tissue, reduce breathing capacity and exacerbate other respiratory problems. *Now* … commonly used surgical masks have three layers of melt-blown textiles made of … polypropylene. We have billions of people putting these microplastics against their mouth, nose and face for hours at a time day after day in the form of masks. How does anyone think that will work out? I mean – what could possibly go wrong? We posted a number of scientific studies on this at davidicke.com, but when I went back to them as I was writing this book the links to the science research website where they were hosted were dead. Anything that challenges the official narrative in any way is either censored or vilified. The official narrative is so unsupportable by the evidence that only deleting the truth can protect it. A study by Chinese scientists still survived – with the usual twist which it why it was still active, I guess. Yes, they found that virtually all the masks they tested increased the daily intake of microplastic fibres, but people should still wear them because the danger from the 'virus' was worse said the crazy 'team' from the Institute of Hydrobiology in Wuhan. Scientists first discovered microplastics in lung tissue of some patients who died of lung cancer

in the 1990s. Subsequent studies have confirmed the potential health damage with the plastic degrading slowly and remaining in the lungs to accumulate in volume. Wuhan researchers used a machine simulating human breathing to establish that masks shed up to nearly 4,000 microplastic fibres in a month with reused masks producing more. Scientists said some masks are laced with toxic chemicals and a variety of compounds seriously restricted for both health and environmental reasons. They include cobalt (used in blue dye) and formaldehyde known to cause watery eyes, burning sensations in the eyes, nose, and throat, plus coughing, wheezing and nausea. No – that must be 'Covid-19'.

Mask 'worms'

There is another and potentially even more sinister content of masks. Mostly new masks of different makes filmed under a microscope around the world have been found to contain strange black fibres or 'worms' that appear to move or 'crawl' by themselves and react to heat and water. The nearest I have seen to them are the self-replicating fibres that are pulled out through the skin of those suffering from Morgellons disease which has been connected to the phenomena of 'chemtrails' which I will bring into the story later on. Morgellons fibres continue to grow outside the body and have a form of artificial intelligence. Black 'worm' fibres in masks have that kind of feel to them and there is a nanotechnology technique called 'worm micelles' which carry and release drugs or anything else you want to deliver to the body. For sure the suppression of humanity by mind altering drugs is the Cult agenda big time and the more excuses they can find to gain access to the body the more opportunities there are to make that happen whether through 'vaccines' or masks pushed against the mouth and nose for hours on end.

So let us summarise the pros and cons of masks:

Against masks: Breathing in your own carbon dioxide; depriving the body and brain of sufficient oxygen; build-up of toxins in the mask that can be breathed into the lungs and cause rashes on the face and 'mask-mouth'; breathing microplastic fibres and toxic chemicals into the lungs; dehumanisation and deleting individualisation by literally making people faceless; destroying human emotional interaction through facial expression and deleting parental connection with their babies which look for guidance to their facial expression.

For masks: They don't protect you from a 'virus' that doesn't exist and even if it did 'viral' particles are so minute they are smaller than the holes in the mask.

Governments, police, supermarkets, businesses, transport companies, and all the rest who seek to impose masks have done no risk assessment on their consequences for health and psychology and are now open to group lawsuits when the impact becomes clear with a cumulative epidemic of respiratory and other disease. Authorities will try to exploit these effects and hide the real cause by dubbing them 'Covid-19'. Can you imagine setting out to force the population to wear health-destroying masks without doing any assessment of the risks? It is criminal and it is evil, but then how many people targeted in this way, who see their children told to wear them all day at school, have asked for a risk assessment? Billions can't be imposed upon by the few unless the billions allow it. Oh, yes, with just a tinge of irony, 85 percent of all masks made worldwide come from *China*.

Wash your hands in toxic shite

'Covid' rules include the use of toxic sanitisers and again the health consequences of constantly applying toxins to be absorbed through the skin is obvious to any level of Renegade Mind. America's Food and Drug Administration (FDA) said that sanitisers are drugs and issued a warning about 75 dangerous brands which contain

methanol used in antifreeze and can cause death, kidney damage and blindness. The FDA circulated the following warning even for those brands that it claims to be safe:

Store hand sanitizer out of the reach of pets and children, and children should use it only with adult supervision. Do not drink hand sanitizer. This is particularly important for young children, especially toddlers, who may be attracted by the pleasant smell or brightly colored bottles of hand sanitizer.

Drinking even a small amount of hand sanitizer can cause alcohol poisoning in children. (However, there is no need to be concerned if your children eat with or lick their hands after using hand sanitizer.) During this coronavirus pandemic, poison control centers have had an increase in calls about accidental ingestion of hand sanitizer, so it is important that adults monitor young children's use.

Do not allow pets to swallow hand sanitizer. If you think your pet has eaten something potentially dangerous, call your veterinarian or a pet poison control center right away. Hand sanitizer is flammable and should be stored away from heat and flames. When using hand sanitizer, rub your hands until they feel completely dry before performing activities that may involve heat, sparks, static electricity, or open flames.

There you go, perfectly safe, then, and that's without even a mention of the toxins absorbed through the skin. Come on kids – sanitise your hands everywhere you go. It will save you from the 'virus'. Put all these elements together of the 'Covid' normal and see how much health and psychology is being cumulatively damaged, even devastated, to 'protect your health'. Makes sense, right? They are only imposing these things because they care, right? *Right?*

Submitting to insanity

Psychological reframing of the population goes very deep and is done in many less obvious ways. I hear people say how contradictory and crazy 'Covid' rules are and how they are ever changing. This is explained away by dismissing those involved as idiots. It is a big mistake. The Cult is delighted if its cold calculation is perceived as incompetence and idiocy when it is anything but. Oh, yes, there are idiots within the system – lots of them – but they are *administering* the Cult agenda, mostly unknowingly. They are not deciding and dictating it. The bulwark against tyranny is self-

respect, always has been, always will be. It is self-respect that has broken every tyranny in history. By its very nature self-respect will not bow to oppression and its perpetrators. There is so little self-respect that it's always the few that overturn dictators. Many may eventually follow, but the few with the iron spines (self-respect) kick it off and generate the momentum. The Cult targets self-respect in the knowledge that once this has gone only submission remains. Crazy, contradictory, ever-changing 'Covid' rules are systematically applied by psychologists to delete self-respect. They *want* you to see that the rules make no sense. It is one thing to decide to do something when *you* have made the choice based on evidence and logic. You still retain your self-respect. It is quite another when you can see what you are being told to do is insane, ridiculous and makes no sense, and *yet you still do it*. Your self-respect is extinguished and this has been happening as ever more obviously stupid and nonsensical things have been demanded and the great majority have complied even when they can see they are stupid and nonsensical.

People walk around in face-nappies knowing they are damaging their health and make no difference to a 'virus'. They do it in fear of not doing it. I know it's daft, but I'll do it anyway. When that happens something dies inside of you and submissive reframing has begun. Next there's a need to hide from yourself that you have conceded your self-respect and you convince yourself that you have not really submitted to fear and intimidation. You begin to believe that you are complying with craziness because it's the right thing to do. When first you concede your self-respect of 2+2 = 4 to 2+2 = 5 you *know* you are compromising your self-respect. Gradually to avoid facing that fact you begin to *believe* that 2+2=5. You have been reframed and I have been watching this process happening in the human psyche on an industrial scale. The Cult is working to break your spirit and one of its major tools in that war is humiliation. I read how former American soldier Bradley Manning (later Chelsea Manning after a sex-change) was treated after being jailed for supplying WikiLeaks with documents exposing the enormity of

government and elite mendacity. Manning was isolated in solitary confinement for eight months, put under 24-hour surveillance, forced to hand over clothing before going to bed, and stand naked for every roll call. This is systematic humiliation. The introduction of anal swab 'Covid' tests in China has been done for the same reason to delete self-respect and induce compliant submission. Anal swabs are mandatory for incoming passengers in parts of China and American diplomats have said they were forced to undergo the indignity which would have been calculated humiliation by the Cult-owned Chinese government that has America in its sights.

Government-people: An abusive relationship

Spirit-breaking psychological techniques include giving people hope and apparent respite from tyranny only to take it away again. This happened in the UK during Christmas, 2020, when the psycho-psychologists and their political lackeys announced an easing of restrictions over the holiday only to reimpose them almost immediately on the basis of yet another lie. There is a big psychological difference between getting used to oppression and being given hope of relief only to have that dashed. Psychologists know this and we have seen the technique used repeatedly. Then there is traumatising people before you introduce more extreme regulations that require compliance. A perfect case was the announcement by the dark and sinister Whitty and Vallance in the UK that 'new data' predicted that 4,000 could die every day over the winter of 2020/2021 if we did not lockdown again. I think they call it lying and after traumatising people with that claim out came Jackboot Johnson the next day with new curbs on human freedom. Psychologists know that a frightened and traumatised mind becomes suggestable to submission and behaviour reframing. Underpinning all this has been to make people fearful and suspicious of each other and see themselves as a potential danger to others. In league with deleted self-respect you have the perfect psychological recipe for self-loathing. The relationship between authority and public is now demonstrably the same as that of

subservience to an abusive partner. These are signs of an abusive relationship explained by psychologist Leslie Becker-Phelps:

Psychological and emotional abuse: Undermining a partner's self-worth with verbal attacks, name-calling, and belittling. Humiliating the partner in public, unjustly accusing them of having an affair, or interrogating them about their every behavior. Keeping partner confused or off balance by saying they were just kidding or blaming the partner for 'making' them act this way … Feigning in public that they care while turning against them in private. This leads to victims frequently feeling confused, incompetent, unworthy, hopeless, and chronically self-doubting. [Apply these techniques to how governments have treated the population since New Year, 2020, and the parallels are obvious.]

Physical abuse: The abuser might physically harm their partner in a range of ways, such as grabbing, hitting, punching, or shoving them. They might throw objects at them or harm them with a weapon. [Observe the physical harm imposed by masks, lockdown, and so on.]

Threats and intimidation: One way abusers keep their partners in line is by instilling fear. They might be verbally threatening, or give threatening looks or gestures. Abusers often make it known that they are tracking their partner's every move. They might destroy their partner's possessions, threaten to harm them, or threaten to harm their family members. Not surprisingly, victims of this abuse often feel anxiety, fear, and panic. [No words necessary.]

Isolation: Abusers often limit their partner's activities, forbidding them to talk or interact with friends or family. They might limit access to a car or even turn off their phone. All of this might be done by physically holding them against their will, but is often accomplished through psychological abuse and intimidation. The more isolated a person feels, the fewer resources they have to help gain perspective on their situation and to escape from it. [No words necessary.]

Economic abuse: Abusers often make their partners beholden to them for money by controlling access to funds of any kind. They might prevent their partner from getting a job or withhold access to money they earn from a job. This creates financial dependency that makes leaving the relationship very difficult. [See destruction of livelihoods and the proposed meagre 'guaranteed income' so long as you do whatever you are told.]

Using children: An abuser might disparage their partner's parenting skills, tell their children lies about their partner, threaten to take custody of their children, or threaten to harm their children. These tactics instil fear and often elicit compliance. [See reframed social service mafia and how children are being mercilessly abused by the state over 'Covid' while their parents look on too frightened to do anything.]

A further recurring trait in an abusive relationship is the abused blaming themselves for their abuse and making excuses for the abuser. We have the public blaming each other for lockdown abuse by government and many making excuses for the government while attacking those who challenge the government. How often we have heard authorities say that rules are being imposed or reimposed only because people have refused to 'behave' and follow the rules. We don't want to do it – it's *you*.

Renegade Minds are an antidote to all of these things. They will never concede their self-respect no matter what the circumstances. Even when apparent humiliation is heaped upon them they laugh in its face and reflect back the humiliation on the abuser where it belongs. Renegade Minds will never wear masks they know are only imposed to humiliate, suppress and damage both physically and psychologically. Consequences will take care of themselves and they will never break their spirit or cause them to concede to tyranny. UK newspaper columnist Peter Hitchens was one of the few in the mainstream media to speak out against lockdowns and forced vaccinations. He then announced he had taken the jab. He wanted to see family members abroad and he believed vaccine passports were inevitable even though they had not yet been introduced. Hitchens

has a questioning and critical mind, but not a Renegade one. If he had no amount of pressure would have made him concede. Hitchens excused his action by saying that the battle has been lost. Renegade Minds never accept defeat when freedom is at stake and even if they are the last one standing the self-respect of not submitting to tyranny is more important than any outcome or any consequence.

That's why Renegade Minds are the only minds that ever changed anything worth changing.

CHAPTER EIGHT

'Reframing' insanity

Insanity is relative. It depends on who has who locked in what cage
Ray Bradbury

'Reframing' a mind means simply to change its perception and behaviour. This can be done subconsciously to such an extent that subjects have no idea they have been 'reframed' while to any observer changes in behaviour and attitudes are obvious.

Human society is being reframed on a ginormous scale since the start of 2020 and here we have the reason why psychologists rather than doctors have been calling the shots. Ask most people who have succumbed to 'Covid' reframing if they have changed and most will say 'no'; but they *have* and fundamentally. The Cult's long-game has been preparing for these times since way back and crucial to that has been to prepare both population and officialdom mentally and emotionally. To use the mind-control parlance they had to reframe the population with a mentality that would submit to fascism and reframe those in government and law enforcement to impose fascism or at least go along with it. The result has been the fact-deleted mindlessness of 'Wokeness' and officialdom that has either enthusiastically or unquestioningly imposed global tyranny demanded by reframed politicians on behalf of psychopathic and deeply evil cultists. 'Cognitive reframing' identifies and challenges the way someone sees the world in the form of situations, experiences and emotions and then restructures those perceptions to view the same set of circumstances in a different way. This can have

benefits if the attitudes are personally destructive while on the other side it has the potential for individual and collective mind control which the subject has no idea has even happened.

Cognitive therapy was developed in the 1960s by Aaron T. Beck who was born in Rhode Island in 1921 as the son of Jewish immigrants from the Ukraine. He became interested in the techniques as a treatment for depression. Beck's daughter Judith S. Beck is prominent in the same field and they founded the Beck Institute for Cognitive Behavior Therapy in Philadelphia in 1994. Cognitive reframing, however, began to be used worldwide by those with a very dark agenda. The Cult reframes politicians to change their attitudes and actions until they are completely at odds with what they once appeared to stand for. The same has been happening to government administrators at all levels, law enforcement, military and the human population. Cultists love mind control for two main reasons: It allows them to control what people think, do and say to secure agenda advancement and, by definition, it calms their legendary insecurity and fear of the unexpected. I have studied mind control since the time I travelled America in 1996. I may have been talking to next to no one in terms of an audience in those years, but my goodness did I gather a phenomenal amount of information and knowledge about so many things including the techniques of mind control. I have described this in detail in other books going back to *The Biggest Secret* in 1998. I met a very large number of people recovering from MKUltra and its offshoots and successors and I began to see how these same techniques were being used on the population in general. This was never more obvious than since the 'Covid' hoax began.

Reframing the enforcers

I have observed over the last two decades and more the very clear transformation in the dynamic between the police, officialdom and the public. I tracked this in the books as the relationship mutated from one of serving the public to seeing them as almost the enemy and certainly a lower caste. There has always been a class divide

based on income and always been some psychopathic, corrupt, and big-I-am police officers. This was different. Wholesale change was unfolding in the collective dynamic; it was less about money and far more about position and perceived power. An us-and-them was emerging. Noses were lifted skyward by government administration and law enforcement and their attitude to the public they were *supposed* to be serving changed to one of increasing contempt, superiority and control. The transformation was so clear and widespread that it had to be planned. Collective attitudes and dynamics do not change naturally and organically that quickly on that scale. I then came across an organisation in Britain called Common Purpose created in the late 1980s by Julia Middleton who would work in the office of Deputy Prime Minister John Prescott during the long and disastrous premiership of war criminal Tony Blair. When Blair speaks the Cult is speaking and the man should have been in jail a long time ago. Common Purpose proclaims itself to be one of the biggest 'leadership development' organisations in the world while functioning as a *charity* with all the financial benefits which come from that. It hosts 'leadership development' courses and programmes all over the world and claims to have 'brought together' what it calls 'leaders' from more than 100 countries on six continents. The modus operandi of Common Purpose can be compared with the work of the UK government's reframing network that includes the Behavioural Insights Team 'nudge unit' and 'Covid' reframing specialists at SPI-B. WikiLeaks described Common Purpose long ago as 'a hidden virus in our government and schools' which is unknown to the general public: 'It recruits and trains "leaders" to be loyal to the directives of Common Purpose and the EU, instead of to their own departments, which they then undermine or subvert, the NHS [National Health Service] being an example.' This is a vital point to understand the 'Covid' hoax. The NHS, and its equivalent around the world, has been utterly reframed in terms of administrators and much of the medical personnel with the transformation underpinned by recruitment policies. The outcome has been the criminal and psychopathic behaviour of the

NHS over 'Covid' and we have seen the same in every other major country. WikiLeaks said Common Purpose trainees are 'learning to rule without regard to democracy' and to usher in a police state (current events explained). Common Purpose operated like a 'glue' and had members in the NHS, BBC, police, legal profession, church, many of Britain's 7,000 quangos, local councils, the Civil Service, government ministries and Parliament, and controlled many RDA's (Regional Development Agencies). Here we have one answer for how and why British institutions and their like in other countries have changed so negatively in relation to the public. This further explains how and why the beyond-disgraceful reframed BBC has become a propaganda arm of 'Covid' fascism. They are all part of a network pursuing the same goal.

By 2019 Common Purpose was quoting a figure of 85,000 'leaders' that had attended its programmes. These 'students' of all ages are known as Common Purpose 'graduates' and they consist of government, state and local government officials and administrators, police chiefs and officers, and a whole range of others operating within the national, local and global establishment. Cressida Dick, Commissioner of the London Metropolitan Police, is the Common Purpose graduate who was the 'Gold Commander' that oversaw what can only be described as the murder of Brazilian electrician Jean Charles de Menezes in 2005. He was held down by psychopathic police and shot seven times in the head by a psychopathic lunatic after being mistaken for a terrorist when he was just a bloke going about his day. Dick authorised officers to pursue and keep surveillance on de Menezes and ordered that he be stopped from entering the underground train system. Police psychopaths took her at her word clearly. She was 'disciplined' for this outrage by being *promoted* – eventually to the top of the 'Met' police where she has been a disaster. Many Chief Constables controlling the police in different parts of the UK are and have been Common Purpose graduates. I have heard the 'graduate' network described as a sort of Mafia or secret society operating within the fabric of government at all levels pursuing a collective policy

ingrained at Common Purpose training events. Founder Julia Middleton herself has said:

> Locally and internationally, Common Purpose graduates will be 'lighting small fires' to create change in their organisations and communities … The Common Purpose effect is best illustrated by the many stories of small changes brought about by leaders, who themselves have changed.

A Common Purpose mission statement declared:

> Common Purpose aims to improve the way society works by expanding the vision, decision-making ability and influence of all kinds of leaders. The organisation runs a variety of educational programmes for leaders of all ages, backgrounds and sectors, in order to provide them with the inspirational, information and opportunities they need to change the world.

Yes, but into what? Since 2020 the answer has become clear.

NLP and the Delphi technique

Common Purpose would seem to be a perfect name or would common programming be better? One of the foundation methods of reaching 'consensus' (group think) is by setting the agenda theme and then encouraging, cajoling or pressuring everyone to agree a 'consensus' in line with the core theme promoted by Common Purpose. The methodology involves the 'Delphi technique', or an adaption of it, in which opinions are expressed that are summarised by a 'facilitator or change agent' at each stage. Participants are 'encouraged' to modify their views in the light of what others have said. Stage by stage the former individual opinions are merged into group consensus which just happens to be what Common Purpose wants them to believe. A key part of this is to marginalise anyone refusing to concede to group think and turn the group against them to apply pressure to conform. We are seeing this very technique used on the general population to make 'Covid' group-thinkers hostile to those who have seen through the bullshit. People can be reframed by using perception manipulation methods such as Neuro-Linguistic Programming (NLP) in which you change perception with the use of

carefully constructed language. An NLP website described the technique this way:

… A method of influencing brain behaviour (the 'neuro' part of the phrase) through the use of language (the 'linguistic' part) and other types of communication to enable a person to 'recode' the way the brain responds to stimuli (that's the 'programming') and manifest new and better behaviours. Neuro-Linguistic Programming often incorporates hypnosis and self-hypnosis to help achieve the change (or 'programming') that is wanted.

British alternative media operation UKColumn has done very detailed research into Common Purpose over a long period. I quoted co-founder and former naval officer Brian Gerrish in my book *Remember Who You Are*, published in 2011, as saying the following years before current times:

It is interesting that many of the mothers who have had children taken by the State speak of the Social Services people being icily cool, emotionless and, as two ladies said in slightly different words, '… like little robots'. We know that NLP is cumulative, so people can be given small imperceptible doses of NLP in a course here, another in a few months, next year etc. In this way, major changes are accrued in their personality, but the day by day change is almost unnoticeable.

In these and other ways 'graduates' have had their perceptions uniformly reframed and they return to their roles in the institutions of government, law enforcement, legal profession, military, 'education', the UK National Health Service and the whole swathe of the establishment structure to pursue a common agenda preparing for the 'post-industrial', 'post-democratic' society. I say 'preparing' but we are now there. 'Post-industrial' is code for the Great Reset and 'post-democratic' is 'Covid' fascism. UKColumn has spoken to partners of those who have attended Common Purpose 'training'. They have described how personalities and attitudes of 'graduates' changed very noticeably for the worse by the time they had completed the course. They had been 'reframed' and told they are the 'leaders' – the special ones – who know better than the population. There has also been the very demonstrable recruitment of psychopaths and narcissists into government administration at all

levels and law enforcement. If you want psychopathy hire psychopaths and you get a simple cause and effect. If you want administrators, police officers and 'leaders' to perceive the public as lesser beings who don't matter then employ narcissists. These personalities are identified using 'psychometrics' that identifies knowledge, abilities, attitudes and personality traits, mostly through carefully-designed questionnaires and tests. As this policy has passed through the decades we have had power-crazy, power-trippers appointed into law enforcement, security and government administration in preparation for current times and the dynamic between public and law enforcement/officialdom has been transformed. UKColumn's Brian Gerrish said of the narcissistic personality:

> Their love of themselves and power automatically means that they will crush others who get in their way. I received a major piece of the puzzle when a friend pointed out that when they made public officials re-apply for their own jobs several years ago they were also required to do psychometric tests. This was undoubtedly the start of the screening process to get 'their' sort of people in post.

How obvious that has been since 2020 although it was clear what was happening long before if people paid attention to the changing public-establishment dynamic.

Change agents

At the centre of events in 'Covid' Britain is the National Health Service (NHS) which has behaved disgracefully in slavishly following the Cult agenda. The NHS management structure is awash with Common Purpose graduates or 'change agents' working to a common cause. Helen Bevan, a Chief of Service Transformation at the NHS Institute for Innovation and Improvement, co-authored a document called 'Towards a million change agents, a review of the social movements literature: implications for large scale change in the NHS'. The document compared a project management approach to that of change and social movements where 'people change

themselves and each other – peer to peer'. Two definitions given for a 'social movement' were:

A group of people who consciously attempt to build a radically new social order; involves people of a broad range of social backgrounds; and deploys politically confrontational and socially disruptive tactics – Cyrus Zirakzadeh 1997

Collective challenges, based on common purposes and social solidarities, in sustained interaction with elites, opponents, and authorities – Sidney Tarrow 1994

Helen Bevan wrote another NHS document in which she defined 'framing' as 'the process by which leaders construct, articulate and put across their message in a powerful and compelling way in order to win people to their cause and call them to action'. I think I could come up with another definition that would be rather more accurate. The National Health Service and institutions of Britain and the wider world have been taken over by reframed 'change agents' and that includes everything from the United Nations to national governments, local councils and social services which have been kidnapping children from loving parents on an extraordinary and gathering scale on the road to the end of parenthood altogether. Children from loving homes are stolen and kidnapped by the state and put into the 'care' (inversion) of the local authority through council homes, foster parents and forced adoption. At the same time children are allowed to be abused without response while many are under council 'care'. UKColumn highlighted the Common Purpose connection between South Yorkshire Police and Rotherham council officers in the case of the scandal in that area of the sexual exploitation of children to which the authorities turned not one blind eye, but both:

We were alarmed to discover that the Chief Executive, the Strategic Director of Children and Young People's Services, the Manager for the Local Strategic Partnership, the Community Cohesion Manager, the Cabinet Member for Cohesion, the Chief Constable and his predecessor had all attended Leadership training courses provided by the pseudo-charity Common Purpose.

Once 'change agents' have secured positions of hire and fire within any organisation things start to move very quickly. Personnel are then hired and fired on the basis of whether they will work towards the agenda the change agent represents. If they do they are rapidly promoted even though they may be incompetent. Those more qualified and skilled who are pre-Common Purpose 'old school' see their careers stall and even disappear. This has been happening for decades in every institution of state, police, 'health' and social services and all of them have been transformed as a result in their attitudes to their jobs and the public. Medical professions, including nursing, which were once vocations for the caring now employ many cold, callous and couldn't give a shit personality types. The UKColumn investigation concluded:

By blurring the boundaries between people, professions, public and private sectors, responsibility and accountability, Common Purpose encourages 'graduates' to believe that as new selected leaders, they can work together, outside of the established political and social structures, to achieve a paradigm shift or CHANGE – so called 'Leading Beyond Authority'. In doing so, the allegiance of the individual becomes 'reframed' on CP colleagues and their NETWORK.

Reframing the Face-Nappies

Nowhere has this process been more obvious than in the police where recruitment of psychopaths and development of unquestioning mind-controlled group-thinkers have transformed law enforcement into a politically-correct 'Woke' joke and a travesty of what should be public service. Today they wear their face-nappies like good little gofers and enforce 'Covid' rules which are fascism under another name. Alongside the specifically-recruited psychopaths we have software minds incapable of free thought. Brian Gerrish again:

An example is the policeman who would not get on a bike for a press photo because he had not done the cycling proficiency course. Normal people say this is political correctness gone mad. Nothing could be further from the truth. The policeman has been reframed, and in his reality it is perfect common sense not to get on the bike 'because he hasn't done the cycling course'.

Another example of this is where the police would not rescue a boy from a pond until they had taken advice from above on the 'risk assessment'. A normal person would have arrived, perhaps thought of the risk for a moment, and dived in. To the police now 'reframed', they followed 'normal' procedure.

There are shocking cases of reframed ambulance crews doing the same. Sheer unthinking stupidity of London Face-Nappies headed by Common Purpose graduate Cressida Dick can be seen in their behaviour at a vigil in March, 2021, for a murdered woman, Sarah Everard. A police officer had been charged with the crime. Anyone with a brain would have left the vigil alone in the circumstances. Instead they 'manhandled' women to stop them breaking 'Covid rules' to betray classic reframing. Minds in the thrall of perception control have no capacity for seeing a situation on its merits and acting accordingly. 'Rules is rules' is their only mind-set. My father used to say that rules and regulations are for the guidance of the intelligent and the blind obedience of the idiot. Most of the intelligent, decent, coppers have gone leaving only the other kind and a few old school for whom the job must be a daily nightmare. The combination of psychopaths and rule-book software minds has been clearly on public display in the 'Covid' era with automaton robots in uniform imposing fascistic 'Covid' regulations on the population without any personal initiative or judging situations on their merits. There are thousands of examples around the world, but I'll make my point with the infamous Derbyshire police in the English East Midlands – the ones who think pouring dye into beauty spots and using drones to track people walking in the countryside away from anyone is called 'policing'. To them there are rules decreed by the government which they have to enforce and in their bewildered state a group gathering in a closed space and someone walking alone in the countryside are the same thing. It is beyond idiocy and enters the realm of clinical insanity.

Police officers in Derbyshire said they were 'horrified' – *horrified* – to find 15 to 20 'irresponsible' kids playing a football match at a closed leisure centre 'in breach of coronavirus restrictions'. When they saw the police the kids ran away leaving their belongings behind and the reframed men and women of Derbyshire police were seeking to establish their identities with a view to fining their parents. The most natural thing for youngsters to do – kicking a ball about – is turned into a criminal activity and enforced by the moronic software programs of Derbyshire police. You find the same mentality in every country. These barely conscious 'horrified' officers said they had to take action because 'we need to ensure these rules are being followed' and 'it is of the utmost importance that you ensure your children are following the rules and regulations for Covid-19'. Had any of them done ten seconds of research to see if this parroting of their masters' script could be supported by any evidence? Nope. Reframed people don't think – others think for them and that's the whole idea of reframing. I have seen police officers one after the other repeating without question word for word what officialdom tells them just as I have seen great swathes of the public doing the same. Ask either for 'their' opinion and out spews what they have been told to think by the official narrative. Police and public may seem to be in different groups, but their mentality is the same. Most people do whatever they are told in fear not doing so or because they believe what officialdom tells them; almost the entirety of the police do what they are told for the same reason. Ultimately it's the tiny inner core of the global Cult that's telling both what to do.

So Derbyshire police were 'horrified'. Oh, really? Why did they think those kids were playing football? It was to relieve the psychological consequences of lockdown and being denied human contact with their friends and interaction, touch and discourse vital to human psychological health. Being denied this month after month has dismantled the psyche of many children and young people as depression and suicide have exploded. Were Derbyshire police *horrified by that*? Are you kidding? Reframed people don't have those

mental and emotional processes that can see how the impact on the psychological health of youngsters is far more dangerous than any 'virus' even if you take the mendacious official figures to be true. The reframed are told (programmed) how to act and so they do. The Derbyshire Chief Constable in the first period of lockdown when the black dye and drones nonsense was going on was Peter Goodman. He was the man who severed the connection between his force and the Derbyshire Constabulary *Male Voice* Choir when he decided that it was not inclusive enough to allow women to join. The fact it was a male voice choir making a particular sound produced by male voices seemed to elude a guy who terrifyingly ran policing in Derbyshire. He retired weeks after his force was condemned as disgraceful by former Supreme Court Justice Jonathan Sumption for their behaviour over extreme lockdown impositions. Goodman was replaced by his deputy Rachel Swann who was in charge when her officers were 'horrified'. The police statement over the boys committing the hanging-offence of playing football included the line about the youngsters being 'irresponsible in the times we are all living through' missing the point that the real relevance of the 'times we are all living through' is the imposition of fascism enforced by psychopaths and reframed minds of police officers playing such a vital part in establishing the fascist tyranny that their own children and grandchildren will have to live in their entire lives. As a definition of insanity that is hard to beat although it might be run close by imposing masks on people that can have a serious effect on their health while wearing a face nappy all day themselves. Once again public and police do it for the same reason – the authorities tell them to and who are they to have the self-respect to say no?

Wokers in uniform

How reframed do you have to be to arrest a *six-year-old* and take him to court for *picking a flower* while waiting for a bus? Brain dead police and officialdom did just that in North Carolina where criminal proceedings happen regularly for children under nine. Attorney Julie Boyer gave the six-year-old crayons and a colouring book

during the 'flower' hearing while the 'adults' decided his fate. County Chief District Court Judge Jay Corpening asked: 'Should a child that believes in Santa Claus, the Easter Bunny and the tooth fairy be making life-altering decisions?' Well, of course not, but common sense has no meaning when you have a common purpose and a reframed mind. Treating children in this way, and police operating in American schools, is all part of the psychological preparation for children to accept a police state as normal all their adult lives. The same goes for all the cameras and biometric tracking technology in schools. Police training is focused on reframing them as snowflake Wokers and this is happening in the military. Pentagon top brass said that 'training sessions on extremism' were needed for troops who asked why they were so focused on the Capitol Building riot when Black Lives Matter riots were ignored. What's the difference between them some apparently and rightly asked. Actually, there is a difference. Five people died in the Capitol riot, only one through violence, and that was a police officer shooting an unarmed protestor. BLM riots killed at least 25 people and cost billions. Asking the question prompted the psychopaths and reframed minds that run the Pentagon to say that more 'education' (programming) was needed. Troop training is all based on psychological programming to make them fodder for the Cult – 'Military men are just dumb, stupid animals to be used as pawns in foreign policy' as Cult-to-his-DNA former Secretary of State Henry Kissinger famously said. Governments see the police in similar terms and it's time for those among them who can see this to defend the people and stop being enforcers of the Cult agenda upon the people.

The US military, like the country itself, is being targeted for destruction through a long list of Woke impositions. Cult-owned gaga 'President' Biden signed an executive order when he took office to allow taxpayer money to pay for transgender surgery for active military personnel and veterans. Are you a man soldier? No, I'm a LGBTQIA+ with a hint of Skoliosexual and Spectrasexual. Oh, good man. Bad choice of words you bigot. The Pentagon announced in March, 2021, the appointment of the first 'diversity and inclusion

officer' for US Special Forces. Richard Torres-Estrada arrived with the publication of a 'D&I Strategic Plan which will guide the enterprise-wide effort to institutionalize and sustain D&I'. If you think a Special Forces 'Strategic Plan' should have something to do with defending America you haven't been paying attention. Defending Woke is now the military's new role. Torres-Estrada has posted images comparing Donald Trump with Adolf Hitler and we can expect no bias from him as a representative of the supposedly non-political Pentagon. Cable news host Tucker Carlson said: 'The Pentagon is now the Yale faculty lounge but with cruise missiles.' Meanwhile Secretary of Defense Lloyd Austin, a board member of weapons-maker Raytheon with stock and compensation interests in October, 2020, worth $1.4 million, said he was purging the military of the 'enemy within' – anyone who isn't Woke and supports Donald Trump. Austin refers to his targets as 'racist extremists' while in true Woke fashion being himself a racist extremist. Pentagon documents pledge to 'eradicate, eliminate and conquer all forms of racism, sexism and homophobia'. The definitions of these are decided by 'diversity and inclusion committees' peopled by those who see racism, sexism and homophobia in every situation and opinion. Woke (the Cult) is dismantling the US military and purging testosterone as China expands its military and gives its troops 'masculinity training'. How do we think that is going to end when this is all Cult coordinated? The US military, like the British military, is controlled by Woke and spineless top brass who just go along with it out of personal career interests.

'Woke' means fast asleep

Mind control and perception manipulation techniques used on individuals to create group-think have been unleashed on the global population in general. As a result many have no capacity to see the obvious fascist agenda being installed all around them or what 'Covid' is really all about. Their brains are firewalled like a computer system not to process certain concepts, thoughts and realisations that are bad for the Cult. The young are most targeted as the adults they

will be when the whole fascist global state is planned to be fully implemented. They need to be prepared for total compliance to eliminate all pushback from entire generations. The Cult has been pouring billions into taking complete control of 'education' from schools to universities via its operatives and corporations and not least Bill Gates as always. The plan has been to transform 'education' institutions into programming centres for the mentality of 'Woke'. James McConnell, professor of psychology at the University of Michigan, wrote in *Psychology Today* in 1970:

> The day has come when we can combine sensory deprivation with drugs, hypnosis, and astute manipulation of reward and punishment, to gain almost absolute control over an individual's behaviour. It should then be possible to achieve a very rapid and highly effective type of brainwashing that would allow us to make dramatic changes in a person's behaviour and personality ...
>
> … We should reshape society so that we all would be trained from birth to want to do what society wants us to do. We have the techniques to do it... no-one owns his own personality you acquired, and there's no reason to believe you should have the right to refuse to acquire a new personality if your old one is anti-social.

This was the potential for mass brainwashing in 1970 and the mentality there displayed captures the arrogant psychopathy that drives it forward. I emphasise that not all young people have succumbed to Woke programming and those that haven't are incredibly impressive people given that today's young are the most perceptually-targeted generations in history with all the technology now involved. Vast swathes of the young generations, however, have fallen into the spell – and that's what it is – of Woke. The Woke mentality and perceptual program is founded on *inversion* and you will appreciate later why that is so significant. Everything with Woke is inverted and the opposite of what it is claimed to be. Woke was a term used in African-American culture from the 1900s and referred to an awareness of social and racial justice. This is not the meaning of the modern version or 'New Woke' as I call it in *The Answer*. Oh, no, Woke today means something very different no matter how much Wokers may seek to hide that and insist Old Woke and New

Woke are the same. See if you find any 'awareness of social justice' here in the modern variety:

- Woke demands 'inclusivity' while excluding anyone with a different opinion and calls for mass censorship to silence other views.

- Woke claims to stand against oppression when imposing oppression is the foundation of all that it does. It is the driver of political correctness which is nothing more than a Cult invention to manipulate the population to silence itself.

- Woke believes itself to be 'liberal' while pursuing a global society that can only be described as fascist (see 'anti-fascist' fascist Antifa).

- Woke calls for 'social justice' while spreading injustice wherever it goes against the common 'enemy' which can be easily identified as a differing view.

- Woke is supposed to be a metaphor for 'awake' when it is solid-gold asleep and deep in a Cult-induced coma that meets the criteria for 'off with the fairies'.

I state these points as obvious facts if people only care to look. I don't do this with a sense of condemnation. We need to appreciate that the onslaught of perceptual programming on the young has been incessant and merciless. I can understand why so many have been reframed, or, given their youth, framed from the start to see the world as the Cult demands. The Cult has had access to their minds day after day in its 'education' system for their entire formative years. Perception is formed from information received and the Cult-created system is a life-long download of information delivered to elicit a particular perception, thus behaviour. The more this has expanded into still new extremes in recent decades and ever-increasing censorship has deleted other opinions and information why wouldn't that lead to a perceptual reframing on a mass scale? I

have described already cradle-to-grave programming and in more recent times the targeting of young minds from birth to adulthood has entered the stratosphere. This has taken the form of skewing what is 'taught' to fit the Cult agenda and the omnipresent techniques of group-think to isolate non-believers and pressure them into line. There has always been a tendency to follow the herd, but we really are in a new world now in relation to that. We have parents who can see the 'Covid' hoax told by their children not to stop them wearing masks at school, being 'Covid' tested or having the 'vaccine' in fear of the peer-pressure consequences of being different. What is 'peer-pressure' if not pressure to conform to group-think? Renegade Minds never group-think and always retain a set of perceptions that are unique to them. Group-think is always underpinned by consequences for not group-thinking. Abuse now aimed at those refusing DNA-manipulating 'Covid vaccines' are a potent example of this. The biggest pressure to conform comes from the very group which is itself being manipulated. 'I am programmed to be part of a hive mind and so you must be.'

Woke control structures in 'education' now apply to every mainstream organisation. Those at the top of the 'education' hierarchy (the Cult) decide the policy. This is imposed on governments through the Cult network; governments impose it on schools, colleges and universities; their leadership impose the policy on teachers and academics and they impose it on children and students. At any level where there is resistance, perhaps from a teacher or university lecturer, they are targeted by the authorities and often fired. Students themselves regularly demand the dismissal of academics (increasingly few) at odds with the narrative that the students have been programmed to believe in. It is quite a thought that students who are being targeted by the Cult become so consumed by programmed group-think that they launch protests and demand the removal of those who are trying to push back against those targeting the students. Such is the scale of perceptual inversion. We see this with 'Covid' programming as the Cult imposes the rules via psycho-psychologists and governments on

shops, transport companies and businesses which impose them on their staff who impose them on their customers who pressure Pushbackers to conform to the will of the Cult which is in the process of destroying them and their families. Scan all aspects of society and you will see the same sequence every time.

Fact free Woke and hijacking the 'left'

There is no more potent example of this than 'Woke', a mentality only made possible by the deletion of factual evidence by an 'education' system seeking to produce an ever more uniform society. Why would you bother with facts when you don't know any? Deletion of credible history both in volume and type is highly relevant. Orwell said: 'Who controls the past controls the future: who controls the present controls the past.' They who control the perception of the past control the perception of the future and they who control the present control the perception of the past through the writing and deleting of history. Why would you oppose the imposition of Marxism in the name of Wokeism when you don't know that Marxism cost at least 100 million lives in the 20th century alone? Watch videos and read reports in which Woker generations are asked basic historical questions – it's mind-blowing. A survey of 2,000 people found that six percent of millennials (born approximately early1980s to early 2000s) believed the Second World War (1939-1945) broke out with the assassination of President Kennedy (in 1963) and one in ten thought Margaret Thatcher was British Prime Minister at the time. She was in office between 1979 and 1990. We are in a post-fact society. Provable facts are no defence against the fascism of political correctness or Silicon Valley censorship. Facts don't matter anymore as we have witnessed with the 'Covid' hoax. Sacrificing uniqueness to the Woke group-think religion is all you are required to do and that means thinking for yourself is the biggest Woke no, no. All religions are an expression of group-think and censorship and Woke is just another religion with an orthodoxy defended by group-think and censorship. Burned at

the stake becomes burned on Twitter which leads back eventually to burned at the stake as Woke humanity regresses to ages past.

The biggest Woke inversion of all is its creators and funders. I grew up in a traditional left of centre political household on a council estate in Leicester in the 1950s and 60s – you know, the left that challenged the power of wealth-hoarding elites and threats to freedom of speech and opinion. In those days students went on marches defending freedom of speech while today's Wokers march for its deletion. What on earth could have happened? Those very elites (collectively the Cult) that we opposed in my youth and early life have funded into existence the antithesis of that former left and hijacked the 'brand' while inverting everything it ever stood for. We have a mentality that calls itself 'liberal' and 'progressive' while acting like fascists. Cult billionaires and their corporations have funded themselves into control of 'education' to ensure that Woke programming is unceasing throughout the formative years of children and young people and that non-Wokers are isolated (that word again) whether they be students, teachers or college professors. The Cult has funded into existence the now colossal global network of Woke organisations that have spawned and promoted all the 'causes' on the Cult wish-list for global transformation and turned Wokers into demanders of them. Does anyone really think it's a coincidence that the Cult agenda for humanity is a carbon (sorry) copy of the societal transformations desired by Woke?? These are only some of them:

Political correctness: The means by which the Cult deletes all public debates that it knows it cannot win if we had the free-flow of information and evidence.

Human-caused 'climate change': The means by which the Cult seeks to transform society into a globally-controlled dictatorship imposing its will over the fine detail of everyone's lives 'to save the planet' which doesn't actually need saving.

Transgender obsession: Preparing collective perception to accept the 'new human' which would not have genders because it would be created technologically and not through procreation. I'll have much more on this in Human 2.0.

Race obsession: The means by which the Cult seeks to divide and rule the population by triggering racial division through the perception that society is more racist than ever when the opposite is the case. Is it perfect in that regard? No. But to compare today with the racism of apartheid and segregation brought to an end by the civil rights movement in the 1960s is to insult the memory of that movement and inspirations like Martin Luther King. Why is the 'anti-racism' industry (which it is) so dominated by privileged white people?

White supremacy: This is a label used by privileged white people to demonise poor and deprived white people pushing back on tyranny to marginalise and destroy them. White people are being especially targeted as the dominant race by number within Western society which the Cult seeks to transform in its image. If you want to change a society you must weaken and undermine its biggest group and once you have done that by using the other groups you next turn on them to do the same … 'Then they came for the Jews and I was not a Jew so I did nothing.'

Mass migration: The mass movement of people from the Middle East, Africa and Asia into Europe, from the south into the United States and from Asia into Australia are another way the Cult seeks to dilute the racial, cultural and political influence of white people on Western society. White people ask why their governments appear to be working against them while being politically and culturally biased towards incoming cultures. Well, here's your answer. In the same way sexually 'straight' people, men and women, ask why the

authorities are biased against them in favour of other sexualities. The answer is the same – that's the way the Cult wants it to be for very sinister motives.

These are all central parts of the Cult agenda and central parts of the Woke agenda and Woke was created and continues to be funded to an immense degree by Cult billionaires and corporations. If anyone begins to say 'coincidence' the syllables should stick in their throat.

Billionaire 'social justice warriors'

Joe Biden is a 100 percent-owned asset of the Cult and the Wokers' man in the White House whenever he can remember his name and for however long he lasts with his rapidly diminishing cognitive function. Even walking up the steps of an aircraft without falling on his arse would appear to be a challenge. He's not an empty-shell puppet or anything. From the minute Biden took office (or the Cult did) he began his executive orders promoting the Woke wish-list. You will see the Woke agenda imposed ever more severely because it's really the *Cult* agenda. Woke organisations and activist networks spawned by the Cult are funded to the extreme so long as they promote what the Cult wants to happen. Woke is funded to promote 'social justice' by billionaires who become billionaires by destroying social justice. The social justice mantra is only a cover for dismantling social justice and funded by billionaires that couldn't give a damn about social justice. Everything makes sense when you see that. One of Woke's premier funders is Cult billionaire financier George Soros who said: 'I am basically there to make money, I cannot and do not look at the social consequences of what I do.' This is the same Soros who has given more than $32 billion to his Open Society Foundations global Woke network and funded Black Lives Matter, mass immigration into Europe and the United States, transgender activism, climate change activism, political correctness and groups targeting 'white supremacy' in the form of privileged white thugs that dominate Antifa. What a scam it all is and when

you are dealing with the unquestioning fact-free zone of Woke scamming them is child's play. All you need to pull it off in all these organisations are a few in-the-know agents of the Cult and an army of naïve, reframed, uninformed, narcissistic, know-nothings convinced of their own self-righteousness, self-purity and virtue.

Soros and fellow billionaires and billionaire corporations have poured hundreds of millions into Black Lives Matter and connected groups and promoted them to a global audience. None of this is motivated by caring about black people. These are the billionaires that have controlled and exploited a system that leaves millions of black people in abject poverty and deprivation which they do absolutely nothing to address. The same Cult networks funding BLM were behind the *slave trade!* Black Lives Matter hijacked a phrase that few would challenge and they have turned this laudable concept into a political weapon to divide society. You know that BLM is a fraud when it claims that *All* Lives Matter, the most inclusive statement of all, is 'racist'. BLM and its Cult masters don't want to end racism. To them it's a means to an end to control all of humanity never mind the colour, creed, culture or background. What has destroying the nuclear family got to do with ending racism? Nothing – but that is one of the goals of BLM and also happens to be a goal of the Cult as I have been exposing in my books for decades. Stealing children from loving parents and giving schools ever more power to override parents is part of that same agenda. BLM is a Marxist organisation and why would that not be the case when the Cult created Marxism *and* BLM? Patrisse Cullors, a BLM co-founder, said in a 2015 video that she and her fellow organisers, including co-founder Alicia Garza, are 'trained Marxists'. The lady known after marriage as Patrisse Khan-Cullors bought a $1.4 million home in 2021 in one of the whitest areas of California with a black population of just 1.6 per cent and has so far bought *four* high-end homes for a total of $3.2 million. How very Marxist. There must be a bit of spare in the BLM coffers, however, when Cult corporations and billionaires have handed over the best part of $100 million. Many black people can see that Black Lives Matter is not

working for them, but against them, and this is still more confirmation. Black journalist Jason Whitlock, who had his account suspended by Twitter for simply linking to the story about the 'Marxist's' home buying spree, said that BLM leaders are 'making millions of dollars off the backs of these dead black men who they wouldn't spit on if they were on fire and alive'.

Black Lies Matter

Cult assets and agencies came together to promote BLM in the wake of the death of career criminal George Floyd who had been jailed a number of times including for forcing his way into the home of a black woman with others in a raid in which a gun was pointed at her stomach. Floyd was filmed being held in a Minneapolis street in 2020 with the knee of a police officer on his neck and he subsequently died. It was an appalling thing for the officer to do, but the same technique has been used by police on peaceful protestors of lockdown without any outcry from the Woke brigade. As unquestioning supporters of the Cult agenda Wokers have supported lockdown and all the 'Covid' claptrap while attacking anyone standing up to the tyranny imposed in its name. Court documents would later include details of an autopsy on Floyd by County Medical Examiner Dr Andrew Baker who concluded that Floyd had taken a fatal level of the drug fentanyl. None of this mattered to fact-free, question-free, Woke. Floyd's death was followed by worldwide protests against police brutality amid calls to defund the police. Throwing babies out with the bathwater is a Woke speciality. In the wake of the murder of British woman Sarah Everard a Green Party member of the House of Lords, Baroness Jones of Moulescoomb (Nincompoopia would have been better), called for a 6pm curfew for all men. This would be in breach of the Geneva Conventions on war crimes which ban collective punishment, but that would never have crossed the black and white Woke mind of Baroness Nincompoopia who would have been far too convinced of her own self-righteousness to compute such details. Many American cities did defund the police in the face of Floyd riots

and after $15 million was deleted from the police budget in Washington DC under useless Woke mayor Muriel Bowser car-jacking alone rose by 300 percent and within six months the US capital recorded its highest murder rate in 15 years. The same happened in Chicago and other cities in line with the Cult/Soros plan to bring fear to streets and neighbourhoods by reducing the police, releasing violent criminals and not prosecuting crime. This is the mob-rule agenda that I have warned in the books was coming for so long. Shootings in the area of Minneapolis where Floyd was arrested increased by 2,500 percent compared with the year before. Defunding the police over George Floyd has led to a big increase in dead people with many of them black. Police protection for politicians making these decisions stayed the same or increased as you would expect from professional hypocrites. The Cult doesn't actually want to abolish the police. It wants to abolish local control over the police and hand it to federal government as the psychopaths advance the Hunger Games Society. Many George Floyd protests turned into violent riots with black stores and businesses destroyed by fire and looting across America fuelled by Black Lives Matter. Woke doesn't do irony. If you want civil rights you must loot the liquor store and the supermarket and make off with a smart TV. It's the only way.

It's not a race war – it's a class war

Black people are patronised by privileged blacks and whites alike and told they are victims of white supremacy. I find it extraordinary to watch privileged blacks supporting the very system and bloodline networks behind the slave trade and parroting the same Cult-serving manipulative crap of their privileged white, often billionaire, associates. It is indeed not a race war but a class war and colour is just a diversion. Black Senator Cory Booker and black Congresswoman Maxine Waters, more residents of Nincompoopia, personify this. Once you tell people they are victims of someone else you devalue both their own responsibility for their plight and the power they have to impact on their reality and experience. Instead

we have: 'You are only in your situation because of whitey – turn on them and everything will change.' It won't change. Nothing changes in our lives unless *we* change it. Crucial to that is never seeing yourself as a victim and always as the creator of your reality. Life is a simple sequence of choice and consequence. Make different choices and you create different consequences. *You* have to make those choices – not Black Lives Matter, the Woke Mafia and anyone else that seeks to dictate your life. Who are they these Wokers, an emotional and psychological road traffic accident, to tell you what to do? Personal empowerment is the last thing the Cult and its Black Lives Matter want black people or anyone else to have. They claim to be defending the underdog while *creating* and perpetuating the underdog. The Cult's worst nightmare is human unity and if they are going to keep blacks, whites and every other race under economic servitude and control then the focus must be diverted from what they have in common to what they can be manipulated to believe divides them. Blacks have to be told that their poverty and plight is the fault of the white bloke living on the street in the same poverty and with the same plight they are experiencing. The difference is that your plight black people is due to him, a white supremacist with 'white privilege' living on the street. Don't unite as one human family against your mutual oppressors and suppressors – fight the oppressor with the white face who is as financially deprived as you are. The Cult knows that as its 'Covid' agenda moves into still new levels of extremism people are going to respond and it has been spreading the seeds of disunity everywhere to stop a united response to the evil that targets *all of us*.

Racist attacks on 'whiteness' are getting ever more outrageous and especially through the American Democratic Party which has an appalling history for anti-black racism. Barack Obama, Joe Biden, Hillary Clinton and Nancy Pelosi all eulogised about Senator Robert Byrd at his funeral in 2010 after a nearly 60-year career in Congress. Byrd was a brutal Ku Klux Klan racist and a violent abuser of Cathy O'Brien in MKUltra. He said he would never fight in the military 'with a negro by my side' and 'rather I should die a thousand times,

and see Old Glory trampled in the dirt never to rise again, than to see this beloved land of ours become degraded by race mongrels, a throwback to the blackest specimen from the wilds'. Biden called Byrd a 'very close friend and mentor'. These 'Woke' hypocrites are not anti-racist they are anti-poor and anti-people not of their perceived class. Here is an illustration of the scale of anti-white racism to which we have now descended. Seriously Woke and moronic *New York Times* contributor Damon Young described whiteness as a 'virus' that 'like other viruses will not die until there are no bodies left for it to infect'. He went on: '… the only way to stop it is to locate it, isolate it, extract it, and kill it.' Young can say that as a black man with no consequences when a white man saying the same in reverse would be facing a jail sentence. *That's* racism. We had super-Woke numbskull senators Tammy Duckworth and Mazie Hirono saying they would object to future Biden Cabinet appointments if he did not nominate more Asian Americans and Pacific Islanders. Never mind the ability of the candidate what do they look like? Duckworth said: 'I will vote for racial minorities and I will vote for LGBTQ, but anyone else I'm not voting for.' Appointing people on the grounds of race is illegal, but that was not a problem for this ludicrous pair. They were on-message and that's a free pass in any situation.

Critical race racism

White children are told at school they are intrinsically racist as they are taught the divisive 'critical race theory'. This claims that the law and legal institutions are inherently racist and that race is a socially constructed concept used by white people to further their economic and political interests at the expense of people of colour. White is a 'virus' as we've seen. Racial inequality results from 'social, economic, and legal differences that white people create between races to maintain white interests which leads to poverty and criminality in minority communities'. I must tell that to the white guy sleeping on the street. The principal of East Side Community School in New York sent white parents a manifesto that called on

them to become 'white traitors' and advocate for full 'white abolition'. These people are teaching your kids when they urgently need a psychiatrist. The 'school' included a chart with 'eight white identities' that ranged from 'white supremacist' to 'white abolition' and defined the behaviour white people must follow to end 'the regime of whiteness'. Woke blacks and their privileged white associates are acting exactly like the slave owners of old and Ku Klux Klan racists like Robert Byrd. They are too full of their own self-purity to see that, but it's true. Racism is not a body type; it's a state of mind that can manifest through any colour, creed or culture.

Another racial fraud is *'equity'*. Not equality of treatment and opportunity – equity. It's a term spun as equality when it means something very different. Equality in its true sense is a raising up while 'equity' is a race to the bottom. Everyone in the same level of poverty is 'equity'. Keep everyone down – that's equity. The Cult doesn't want anyone in the human family to be empowered and BLM leaders, like all these 'anti-racist' organisations, continue their privileged, pampered existence by perpetuating the perception of gathering racism. When is the last time you heard an 'anti-racist' or 'anti-Semitism' organisation say that acts of racism and discrimination have *fallen?* It's not in the interests of their fund-raising and power to influence and the same goes for the professional soccer anti-racism operation, Kick It Out. Two things confirmed that the Black Lives Matter riots in the summer of 2020 were Cult creations. One was that while anti-lockdown protests were condemned in this same period for 'transmitting 'Covid' the authorities supported mass gatherings of Black Lives Matter supporters. I even saw self-deluding people claiming to be doctors say the two types of protest were not the same. No – the non-existent 'Covid' was in favour of lockdowns and attacked those that protested against them while 'Covid' supported Black Lives Matter and kept well away from its protests. The whole thing was a joke and as lockdown protestors were arrested, often brutally, by reframed Face-Nappies we had the grotesque sight of police officers taking the knee to Black Lives Matter, a Cult-funded Marxist

organisation that supports violent riots and wants to destroy the nuclear family and white people.

He's not white? Shucks!

Woke obsession with race was on display again when ten people were shot dead in Boulder, Colorado, in March, 2021. Cult-owned Woke TV channels like CNN said the shooter appeared to be a white man and Wokers were on Twitter condemning 'violent white men' with the usual mantras. Then the shooter's name was released as Ahmad Al Aliwi Alissa, an anti-Trump Arab-American, and the sigh of disappointment could be heard five miles away. Never mind that ten people were dead and what that meant for their families. Race baiting was all that mattered to these sick Cult-serving people like Barack Obama who exploited the deaths to further divide America on racial grounds which is his job for the Cult. This is the man that 'racist' white Americans made the first black president of the United States and then gave him a second term. Not-very-bright Obama has become filthy rich on the back of that and today appears to have a big influence on the Biden administration. Even so he's still a downtrodden black man and a victim of white supremacy. This disingenuous fraud reveals the contempt he has for black people when he puts on a Deep South Alabama accent whenever he talks to them, no, *at* them.

Another BLM red flag was how the now fully-Woke (fully-Cult) and fully-virtue-signalled professional soccer authorities had their teams taking the knee before every match in support of Marxist Black Lives Matter. Soccer authorities and clubs displayed 'Black Lives Matter' on the players' shirts and flashed the name on electronic billboards around the pitch. Any fans that condemned what is a Freemasonic taking-the-knee ritual were widely condemned as you would expect from the Woke virtue-signallers of professional sport and the now fully-Woke media. We have reverse racism in which you are banned from criticising any race or culture except for white people for whom anything goes – say what you like, no problem. What has this got to do with racial harmony and

equality? We've had black supremacists from Black Lives Matter telling white people to fall to their knees in the street and apologise for their white supremacy. Black supremacists acting like white supremacist slave owners of the past couldn't breach their self-obsessed, race-obsessed sense of self-purity. Joe Biden appointed a race-obsessed black supremacist Kristen Clarke to head the Justice Department Civil Rights Division. Clarke claimed that blacks are endowed with 'greater mental, physical and spiritual abilities' than whites. If anyone reversed that statement they would be vilified. Clarke is on-message so no problem. She's never seen a black-white situation in which the black figure is anything but a virtuous victim and she heads the Civil Rights Division which should treat everyone the same or it isn't civil rights. Another perception of the Renegade Mind: If something or someone is part of the Cult agenda they will be supported by Woke governments and media no matter what. If they're not, they will be condemned and censored. It really is that simple and so racist Clarke prospers despite (make that because of) her racism.

The end of culture

Biden's administration is full of such racial, cultural and economic bias as the Cult requires the human family to be divided into warring factions. We are now seeing racially-segregated graduations and everything, but everything, is defined through the lens of perceived 'racism. We have 'racist' mathematics, 'racist' food and even 'racist' *plants*. World famous Kew Gardens in London said it was changing labels on plants and flowers to tell its pre-'Covid' more than two million visitors a year how racist they are. Kew director Richard Deverell said this was part of an effort to 'move quickly to decolonise collections' after they were approached by one Ajay Chhabra 'an actor with an insight into how sugar cane was linked to slavery'. They are *plants* you idiots. 'Decolonisation' in the Woke manual really means colonisation of society with its mentality and by extension colonisation by the Cult. We are witnessing a new Chinese-style 'Cultural Revolution' so essential to the success of all

Marxist takeovers. Our cultural past and traditions have to be swept away to allow a new culture to be built-back-better. Woke targeting of long-standing Western cultural pillars including historical monuments and cancelling of historical figures is what happened in the Mao revolution in China which 'purged remnants of capitalist and traditional elements from Chinese society' and installed Maoism as the dominant ideology'. For China see the Western world today and for 'dominant ideology' see Woke. Better still see Marxism or Maoism. The 'Covid' hoax has specifically sought to destroy the arts and all elements of Western culture from people meeting in a pub or restaurant to closing theatres, music venues, sports stadiums, places of worship and even banning *singing*. Destruction of Western society is also why criticism of any religion is banned except for Christianity which again is the dominant religion as white is the numerically-dominant race. Christianity may be fading rapidly, but its history and traditions are weaved through the fabric of Western society. Delete the pillars and other structures will follow until the whole thing collapses. I am not a Christian defending that religion when I say that. I have no religion. It's just a fact. To this end Christianity has itself been turned Woke to usher its own downfall and its ranks are awash with 'change agents' – knowing and unknowing – at every level including Pope Francis (*definitely* knowing) and the clueless Archbishop of Canterbury Justin Welby (possibly not, but who can be sure?). Woke seeks to coordinate attacks on Western culture, traditions, and ways of life through 'intersectionality' defined as 'the complex, cumulative way in which the effects of multiple forms of discrimination (such as racism, sexism, and classism) combine, overlap, or intersect especially in the experiences of marginalised individuals or groups'. Wade through the Orwellian Woke-speak and this means coordinating disparate groups in a common cause to overthrow freedom and liberal values.

The entire structure of public institutions has been infested with Woke – government at all levels, political parties, police, military, schools, universities, advertising, media and trade unions. This abomination has been achieved through the Cult web by appointing

Wokers to positions of power and battering non-Wokers into line through intimidation, isolation and threats to their job. Many have been fired in the wake of the empathy-deleted, vicious hostility of 'social justice' Wokers and the desire of gutless, spineless employers to virtue-signal their Wokeness. Corporations are filled with Wokers today, most notably those in Silicon Valley. Ironically at the top they are not Woke at all. They are only exploiting the mentality their Cult masters have created and funded to censor and enslave while the Wokers cheer them on until it's their turn. Thus the Woke 'liberal left' is an inversion of the traditional liberal left. Campaigning for justice on the grounds of power and wealth distribution has been replaced by campaigning for identity politics. The genuine traditional left would never have taken money from today's billionaire abusers of fairness and justice and nor would the billionaires have wanted to fund that genuine left. It would not have been in their interests to do so. The division of opinion in those days was between the haves and have nots. This all changed with Cult manipulated and funded identity politics. The division of opinion today is between Wokers and non-Wokers and not income brackets. Cult corporations and their billionaires may have taken wealth disparity to cataclysmic levels of injustice, but as long as they speak the language of Woke, hand out the dosh to the Woke network and censor the enemy they are 'one of us'. Billionaires who don't give a damn about injustice are laughing at them till their bellies hurt. Wokers are not even close to self-aware enough to see that. The transformed 'left' dynamic means that Wokers who drone on about 'social justice' are funded by billionaires that have destroyed social justice the world over. It's *why* they are billionaires.

The climate con

Nothing encapsulates what I have said more comprehensively than the hoax of human-caused global warming. I have detailed in my books over the years how Cult operatives and organisations were the pump-primers from the start of the climate con. A purpose-built vehicle for this is the Club of Rome established by the Cult in 1968

with the Rockefellers and Rothschilds centrally involved all along. Their gofer frontman Maurice Strong, a Canadian oil millionaire, hosted the Earth Summit in Rio de Janeiro, Brazil, in 1992 where the global 'green movement' really expanded in earnest under the guiding hand of the Cult. The Earth Summit established Agenda 21 through the Cult-created-and-owned United Nations to use the illusion of human-caused climate change to justify the transformation of global society to save the world from climate disaster. It is a No-Problem-Reaction-Solution sold through governments, media, schools and universities as whole generations have been terrified into believing that the world was going to end in their lifetimes unless what old people had inflicted upon them was stopped by a complete restructuring of how everything is done. Chill, kids, it's all a hoax. Such restructuring is precisely what the Cult agenda demands (purely by coincidence of course). Today this has been given the codename of the Great Reset which is only an updated term for Agenda 21 and its associated Agenda 2030. The latter, too, is administered through the UN and was voted into being by the General Assembly in 2015. Both 21 and 2030 seek centralised control of all resources and food right down to the raindrops falling on your own land. These are some of the demands of Agenda 21 established in 1992. See if you recognise this society emerging today:

- End national sovereignty
- State planning and management of all land resources, ecosystems, deserts, forests, mountains, oceans and fresh water; agriculture; rural development; biotechnology; and ensuring *'equity'*
- The state to 'define the role' of business and financial resources
- Abolition of private property
- 'Restructuring' the family unit (see BLM)
- Children raised by the state
- People told what their job will be
- Major restrictions on movement
- Creation of 'human settlement zones'

- Mass resettlement as people are forced to vacate land where they live
- Dumbing down education
- Mass global depopulation in pursuit of all the above

The United Nations was created as a Trojan horse for world government. With the climate con of critical importance to promoting that outcome you would expect the UN to be involved. Oh, it's involved all right. The UN is promoting Agenda 21 and Agenda 2030 justified by 'climate change' while also driving the climate hoax through its Intergovernmental Panel on Climate Change (IPCC), one of the world's most corrupt organisations. The IPCC has been lying ferociously and constantly since the day it opened its doors with the global media hanging unquestioningly on its every mendacious word. The Green movement is entirely Woke and has long lost its original environmental focus since it was co-opted by the Cult. An obsession with 'global warming' has deleted its values and scrambled its head. I experienced a small example of what I mean on a beautiful country walk that I have enjoyed several times a week for many years. The path merged into the fields and forests and you felt at one with the natural world. Then a 'Green' organisation, the Hampshire and Isle of Wight Wildlife Trust, took over part of the land and proceeded to cut down a large number of trees, including mature ones, to install a horrible big, bright steel 'this-is-ours-stay-out' fence that destroyed the whole atmosphere of this beautiful place. No one with a feel for nature would do that. Day after day I walked to the sound of chainsaws and a magnificent mature weeping willow tree that I so admired was cut down at the base of the trunk. When I challenged a Woke young girl in a green shirt (of course) about this vandalism she replied: 'It's a weeping willow – it will grow back.' This is what people are paying for when they donate to the Hampshire and Isle of Wight Wildlife Trust and many other 'green' organisations today. It is not the environmental movement that I knew and instead has become a support-system – as with Extinction Rebellion – for a very dark agenda.

Private jets for climate justice

The Cult-owned, Gates-funded, World Economic Forum and its founder Klaus Schwab were behind the emergence of Greta Thunberg to harness the young behind the climate agenda and she was invited to speak to the world at … the UN. Schwab published a book, *Covid-19: The Great Reset* in 2020 in which he used the 'Covid' hoax and the climate hoax to lay out a new society straight out of Agenda 21 and Agenda 2030. Bill Gates followed in early 2021 when he took time out from destroying the world to produce a book in his name about the way to save it. Gates flies across the world in private jets and admitted that 'I probably have one of the highest greenhouse gas footprints of anyone on the planet … my personal flying alone is gigantic.' He has also bid for the planet's biggest private jet operator. Other climate change saviours who fly in private jets include John Kerry, the US Special Presidential Envoy for Climate, and actor Leonardo DiCaprio, a 'UN Messenger of Peace with special focus on climate change'. These people are so full of bullshit they could corner the market in manure. We mustn't be sceptical, though, because the Gates book, *How to Avoid a Climate Disaster: The Solutions We Have and the Breakthroughs We Need*, is a genuine attempt to protect the world and not an obvious pile of excrement attributed to a mega-psychopath aimed at selling his masters' plans for humanity. The Gates book and the other shite-pile by Klaus Schwab could have been written by the same person and may well have been. Both use 'climate change' and 'Covid' as the excuses for their new society and by coincidence the Cult's World Economic Forum and Bill and Melinda Gates Foundation promote the climate hoax and hosted Event 201 which pre-empted with a 'simulation' the very 'coronavirus' hoax that would be simulated for real on humanity within weeks. The British 'royal' family is promoting the 'Reset' as you would expect through Prince 'climate change caused the war in Syria' Charles and his hapless son Prince William who said that we must 'reset our relationship with nature and our trajectory as a species' to avoid a climate disaster. Amazing how many promotors of the 'Covid' and 'climate change' control

systems are connected to Gates and the World Economic Forum. A 'study' in early 2021 claimed that carbon dioxide emissions must fall by the equivalent of a global lockdown roughly every two years for the next decade to save the planet. The 'study' appeared in the same period that the Schwab mob claimed in a video that lockdowns destroying the lives of billions are good because they make the earth 'quieter' with less 'ambient noise'. They took down the video amid a public backlash for such arrogant, empathy-deleted stupidity You see, however, where they are going with this. Corinne Le Quéré, a professor at the Tyndall Centre for Climate Change Research, University of East Anglia, was lead author of the climate lockdown study, and she writes for … the World Economic Forum. Gates calls in 'his' book for changing 'every aspect of the economy' (long-time Cult agenda) and for humans to eat synthetic 'meat' (predicted in my books) while cows and other farm animals are eliminated. Australian TV host and commentator Alan Jones described what carbon emission targets would mean for farm animals in Australia alone if emissions were reduced as demanded by 35 percent by 2030 and zero by 2050:

Well, let's take agriculture, the total emissions from agriculture are about 75 million tonnes of carbon dioxide, equivalent. Now reduce that by 35 percent and you have to come down to 50 million tonnes, I've done the maths. So if you take for example 1.5 million cows, you're going to have to reduce the herd by 525,000 [by] 2030, nine years, that's 58,000 cows a year. The beef herd's 30 million, reduce that by 35 percent, that's 10.5 million, which means 1.2 million cattle have to go every year between now and 2030. This is insanity!

There are 75 million sheep. Reduce that by 35 percent, that's 26 million sheep, that's almost 3 million a year. So under the Paris Agreement over 30 million beasts. dairy cows, cattle, pigs and sheep would go. More than 8,000 every minute of every hour for the next decade, do these people know what they're talking about?

Clearly they don't at the level of campaigners, politicians and administrators. The Cult *does* know; that's the outcome it wants. We are faced with not just a war on humanity. Animals and the natural world are being targeted and I have been saying since the 'Covid' hoax began that the plan eventually was to claim that the 'deadly virus' is able to jump from animals, including farm animals and

domestic pets, to humans. Just before this book went into production came this story: 'Russia registers world's first Covid-19 vaccine for cats & dogs as makers of Sputnik V warn pets & farm animals could spread virus'. The report said 'top scientists warned that the deadly pathogen could soon begin spreading through homes and farms' and 'the next stage is the infection of farm and domestic animals'. Know the outcome and you'll see the journey. Think what that would mean for animals and keep your eye on a term called zoonosis or zoonotic diseases which transmit between animals and humans. The Cult wants to break the connection between animals and people as it does between people and people. Farm animals fit with the Cult agenda to transform food from natural to synthetic.

The gas of life is killing us

There can be few greater examples of Cult inversion than the condemnation of carbon dioxide as a dangerous pollutant when it is the gas of life. Without it the natural world would be dead and so we would all be dead. We breathe in oxygen and breathe out carbon dioxide while plants produce oxygen and absorb carbon dioxide. It is a perfect symbiotic relationship that the Cult wants to dismantle for reasons I will come to in the final two chapters. Gates, Schwab, other Cult operatives and mindless repeaters, want the world to be 'carbon neutral' by at least 2050 and the earlier the better. 'Zero carbon' is the cry echoed by lunatics calling for 'Zero Covid' when we already have it. These carbon emission targets will deindustrialise the world in accordance with Cult plans – the post-industrial, post-democratic society – and with so-called renewables like solar and wind not coming even close to meeting human energy needs blackouts and cold are inevitable. Texans got the picture in the winter of 2021 when a snow storm stopped wind turbines and solar panels from working and the lights went down along with water which relies on electricity for its supply system. Gates wants everything to be powered by electricity to ensure that his masters have the kill switch to stop all human activity, movement, cooking, water and warmth any time they like. The climate lie is so

stupendously inverted that it claims we must urgently reduce carbon dioxide when we *don't have enough*.

Co2 in the atmosphere is a little above 400 parts per million when the optimum for plant growth is 2,000 ppm and when it falls anywhere near 150 ppm the natural world starts to die and so do we. It fell to as low as 280 ppm in an 1880 measurement in Hawaii and rose to 413 ppm in 2019 with industrialisation which is why the planet has become *greener* in the industrial period. How insane then that psychopathic madman Gates is not satisfied only with blocking the rise of Co2. He's funding technology to suck it out of the atmosphere. The reason why will become clear. The industrial era is not destroying the world through Co2 and has instead turned around a potentially disastrous ongoing fall in Co2. Greenpeace co-founder and scientist Patrick Moore walked away from Greenpeace in 1986 and has exposed the green movement for fear-mongering and lies. He said that 500 million years ago there was *17 times* more Co2 in the atmosphere than we have today and levels have been falling for hundreds of millions of years. In the last 150 million years Co2 levels in Earth's atmosphere had reduced by *90 percent*. Moore said that by the time humanity began to unlock carbon dioxide from fossil fuels we were at '38 seconds to midnight' and in that sense: 'Humans are [the Earth's] salvation.' Moore made the point that only half the Co2 emitted by fossil fuels stays in the atmosphere and we should remember that all pollution pouring from chimneys that we are told is carbon dioxide is in fact nothing of the kind. It's pollution. Carbon dioxide is an invisible gas.

William Happer, Professor of Physics at Princeton University and long-time government adviser on climate, has emphasised the Co2 deficiency for maximum growth and food production. Greenhouse growers don't add carbon dioxide for a bit of fun. He said that most of the warming in the last 100 years, after the earth emerged from the super-cold period of the 'Little Ice Age' into a natural warming cycle, was over by 1940. Happer said that a peak year for warming in 1988 can be explained by a 'monster El Nino' which is a natural and cyclical warming of the Pacific that has nothing to do with 'climate

change'. He said the effect of Co2 could be compared to painting a wall with red paint in that once two or three coats have been applied it didn't matter how much more you slapped on because the wall will not get much redder. Almost all the effect of the rise in Co2 has already happened, he said, and the volume in the atmosphere would now have to *double* to increase temperature by a single degree. Climate hoaxers know this and they have invented the most ridiculously complicated series of 'feedback' loops to try to overcome this rather devastating fact. You hear puppet Greta going on cluelessly about feedback loops and this is why.

The Sun affects temperature? No you *climate denier*

Some other nonsense to contemplate: Climate graphs show that rises in temperature do not follow rises in Co2 – *it's the other way round* with a lag between the two of some 800 years. If we go back 800 years from present time we hit the Medieval Warm Period when temperatures were higher than now without any industrialisation and this was followed by the Little Ice Age when temperatures plummeted. The world was still emerging from these centuries of serious cold when many climate records began which makes the ever-repeated line of the 'hottest year since records began' meaningless when you are not comparing like with like. The coldest period of the Little Ice Age corresponded with the lowest period of sunspot activity when the Sun was at its least active. Proper scientists will not be at all surprised by this when it confirms the obvious fact that earth temperature is affected by the scale of Sun activity and the energetic power that it subsequently emits; but when is the last time you heard a climate hoaxer talking about the Sun as a source of earth temperature?? Everything has to be focussed on Co2 which makes up just 0.117 percent of so-called greenhouse gases and only a fraction of even that is generated by human activity. The rest is natural. More than *90 percent* of those greenhouse gases are water vapour and clouds (Fig 9). Ban moisture I say. Have you noticed that the climate hoaxers no longer use the polar bear as their promotion image? That's because far from becoming extinct polar

bear communities are stable or thriving. Joe Bastardi, American meteorologist, weather forecaster and outspoken critic of the climate lie, documents in his book *The Climate Chronicles* how weather patterns and events claimed to be evidence of climate change have been happening since long before industrialisation: 'What happened before naturally is happening again, as is to be expected given the cyclical nature of the climate due to the design of the planet.' If you read the detailed background to the climate hoax in my other books you will shake your head and wonder how anyone could believe the crap which has spawned a multi-trillion dollar industry based on absolute garbage (see HIV causes AIDs and Sars-Cov-2 causes 'Covid-19'). Climate and 'Covid' have much in common given they have the same source. They both have the contradictory *everything* factor in which everything is explained by reference to them. It's hot – 'it's climate change'. It's cold – 'it's climate change'. I got a sniffle – 'it's Covid'. I haven't got a sniffle – 'it's Covid'. Not having a sniffle has to be a symptom of 'Covid'. Everything is and not having a sniffle is especially dangerous if you are a slow walker. For sheer audacity I offer you a Cambridge University 'study' that actually linked 'Covid' to 'climate change'. It had to happen eventually. They concluded that climate change played a role in 'Covid-19' spreading from animals to humans because … wait for it … I kid you not … *the two groups were forced closer together as populations grow*. Er, that's it. The whole foundation on which this depended was that 'Bats are the likely zoonotic origin of SARS-CoV-1 and SARS-CoV-2'. Well, they are not. They are nothing to do with it. Apart from bats not being the origin and therefore 'climate change' effects on bats being irrelevant I am in awe of their academic insight. Where would we be without them? Not where we are that's for sure.

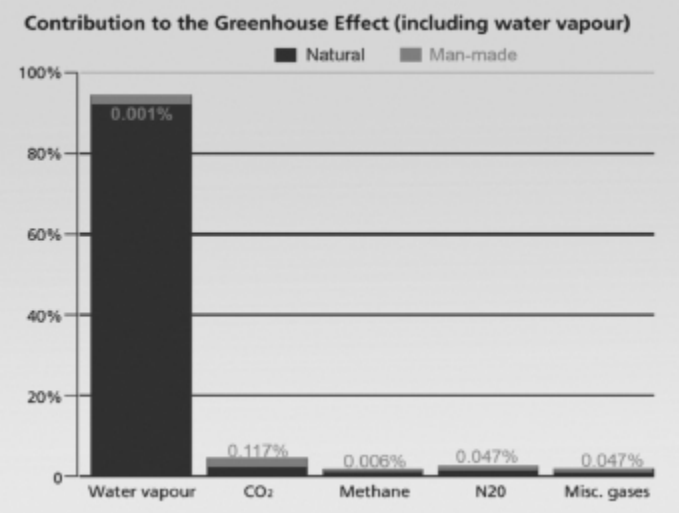


Figure 9: The idea that the gas of life is disastrously changing the climate is an insult to brain cell activity.

One other point about the weather is that climate modification is now well advanced and not every major weather event is natural – or earthquake come to that. I cover this subject at some length in other books. China is openly planning a rapid expansion of its weather modification programme which includes changing the climate in an area more than one and a half times the size of India. China used weather manipulation to ensure clear skies during the 2008 Olympics in Beijing. I have quoted from US military documents detailing how to employ weather manipulation as a weapon of war and they did that in the 1960s and 70s during the conflict in Vietnam with Operation Popeye manipulating monsoon rains for military purposes. Why would there be international treaties on weather modification if it wasn't possible? Of course it is. Weather is energetic information and it can be changed.

How was the climate hoax pulled off? See 'Covid'

If you can get billions to believe in a 'virus' that doesn't exist you can get them to believe in human-caused climate change that doesn't exist. Both are being used by the Cult to transform global society in the way it has long planned. Both hoaxes have been achieved in pretty much the same way. First you declare a lie is a fact. There's a

'virus' you call SARS-Cov-2 or humans are warming the planet with their behaviour. Next this becomes, via Cult networks, the foundation of government, academic and science policy and belief. Those who parrot the mantra are given big grants to produce research that confirms the narrative is true and ever more 'symptoms' are added to make the 'virus'/'climate change' sound even more scary. Scientists and researchers who challenge the narrative have their grants withdrawn and their careers destroyed. The media promote the lie as the unquestionable truth and censor those with an alternative view or evidence. A great percentage of the population believe what they are told as the lie becomes an everybody-knows-that and the believing-masses turn on those with a mind of their own. The technique has been used endlessly throughout human history. Wokers are the biggest promotors of the climate lie *and* 'Covid' fascism because their minds are owned by the Cult; their sense of self-righteous self-purity knows no bounds; and they exist in a bubble of reality in which facts are irrelevant and only get in the way of looking without seeing.

Running through all of this like veins in a blue cheese is control of information, which means control of perception, which means control of behaviour, which collectively means control of human society. The Cult owns the global media and Silicon Valley fascists for the simple reason that it *has* to. Without control of information it can't control perception and through that human society. Examine every facet of the Cult agenda and you will see that anything supporting its introduction is never censored while anything pushing back is always censored. I say again: Psychopaths that know why they are doing this must go before Nuremberg trials and those that follow their orders must trot along behind them into the same dock. 'I was just following orders' didn't work the first time and it must not work now. Nuremberg trials must be held all over the world before public juries for politicians, government officials, police, compliant doctors, scientists and virologists, and all Cult operatives such as Gates, Tedros, Fauci, Vallance, Whitty, Ferguson, Zuckerberg, Wojcicki, Brin, Page, Dorsey, the whole damn lot of

them – including, no *especially*, the psychopath psychologists. Without them and the brainless, gutless excuses for journalists that have repeated their lies, none of this could be happening. Nobody can be allowed to escape justice for the psychological and economic Armageddon they are all responsible for visiting upon the human race.

As for the compliant, unquestioning, swathes of humanity, and the self-obsessed, all-knowing ignorance of the Wokers … don't start me. God help their kids. God help their grandkids. God *help them*.

CHAPTER NINE

We must have it? So what is it?

Well I won't back down. No, I won't back down. You can stand me up at the Gates of Hell. But I won't back down
Tom Petty

I will now focus on the genetically-manipulating 'Covid vaccines' which do not meet this official definition of a vaccine by the US Centers for Disease Control (CDC): 'A product that stimulates a person's immune system to produce immunity to a specific disease, protecting the person from that disease.' On that basis 'Covid vaccines' are not a vaccine in that the makers don't even claim they stop infection or transmission.

They are instead part of a multi-levelled conspiracy to change the nature of the human body and what it means to be 'human' and to depopulate an enormous swathe of humanity. What I shall call Human 1.0 is on the cusp of becoming Human 2.0 and for very sinister reasons. Before I get to the 'Covid vaccine' in detail here's some background to vaccines in general. Government regulators do not test vaccines – the makers do – and the makers control which data is revealed and which isn't. Children in America are given 50 vaccine doses by age six and 69 by age 19 and the effect of the whole combined schedule has never been tested. Autoimmune diseases when the immune system attacks its own body have soared in the mass vaccine era and so has disease in general in children and the young. Why wouldn't this be the case when vaccines target the *immune system?* The US government gave Big Pharma drug

companies immunity from prosecution for vaccine death and injury in the 1986 National Childhood Vaccine Injury Act (NCVIA) and since then the government (taxpayer) has been funding compensation for the consequences of Big Pharma vaccines. The criminal and satanic drug giants can't lose and the vaccine schedule has increased dramatically since 1986 for this reason. There is no incentive to make vaccines safe and a big incentive to make money by introducing ever more. Even against a ridiculously high bar to prove vaccine liability, and with the government controlling the hearing in which it is being challenged for compensation, the vaccine court has so far paid out more than $4 billion. These are the vaccines we are told are safe and psychopaths like Zuckerberg censor posts saying otherwise. The immunity law was even justified by a ruling that vaccines by their nature were 'unavoidably unsafe'.

Check out the ingredients of vaccines and you will be shocked if you are new to this. *They put that in children's bodies?? What??* Try aluminium, a brain toxin connected to dementia, aborted foetal tissue and formaldehyde which is used to embalm corpses. World-renowned aluminium expert Christopher Exley had his research into the health effect of aluminium in vaccines shut down by Keele University in the UK when it began taking funding from the Bill and Melinda Gates Foundation. Research when diseases 'eradicated' by vaccines began to decline and you will find the fall began long *before* the vaccine was introduced. Sometimes the fall even plateaued after the vaccine. Diseases like scarlet fever for which there was no vaccine declined in the same way because of environmental and other factors. A perfect case in point is the polio vaccine. Polio began when lead arsenate was first sprayed as an insecticide and residues remained in food products. Spraying started in 1892 and the first US polio epidemic came in Vermont in 1894. The simple answer was to stop spraying, but Rockefeller-created Big Pharma had a better idea. Polio was decreed to be caused by the *poliovirus* which 'spreads from person to person and can infect a person's spinal cord'. Lead arsenate was replaced by the lethal DDT which had the same effect of causing paralysis by damaging the brain and central nervous

system. Polio plummeted when DDT was reduced and then banned, but the vaccine is still given the credit for something it didn't do. Today by far the biggest cause of polio is the vaccines promoted by Bill Gates. Vaccine justice campaigner Robert Kennedy Jr, son of assassinated (by the Cult) US Attorney General Robert Kennedy, wrote:

> In 2017, the World Health Organization (WHO) reluctantly admitted that the global explosion in polio is predominantly vaccine strain. The most frightening epidemics in Congo, Afghanistan, and the Philippines, are all linked to vaccines. In fact, by 2018, 70% of global polio cases were vaccine strain.

Vaccines make fortunes for Cult-owned Gates and Big Pharma while undermining the health and immune systems of the population. We had a glimpse of the mentality behind the Big Pharma cartel with a report on WION (World is One News), an international English language TV station based in India, which exposed the extraordinary behaviour of US drug company Pfizer over its 'Covid vaccine'. The WION report told how Pfizer had made fantastic demands of Argentina, Brazil and other countries in return for its 'vaccine'. These included immunity from prosecution, even for Pfizer negligence, government insurance to protect Pfizer from law suits and handing over as collateral sovereign assets of the country to include Argentina's bank reserves, military bases and embassy buildings. Pfizer demanded the same of Brazil in the form of waiving sovereignty of its assets abroad; exempting Pfizer from Brazilian laws; and giving Pfizer immunity from all civil liability. This is a 'vaccine' developed with government funding. Big Pharma is evil incarnate as a creation of the Cult and all must be handed tickets to Nuremberg.

Phantom 'vaccine' for a phantom 'disease'

I'll expose the 'Covid vaccine' fraud and then go on to the wider background of why the Cult has set out to 'vaccinate' every man, woman and child on the planet for an alleged 'new disease' with a survival rate of 99.77 percent (or more) even by the grotesquely-

manipulated figures of the World Health Organization and Johns Hopkins University. The 'infection' to 'death' ratio is 0.23 to 0.15 percent according to Stanford epidemiologist Dr John Ioannidis and while estimates vary the danger remains tiny. I say that if the truth be told the fake infection to fake death ratio is zero. Never mind all the evidence I have presented here and in *The Answer* that there is no 'virus' let us just focus for a moment on that death-rate figure of say 0.23 percent. The figure includes all those worldwide who have tested positive with a test not testing for the 'virus' and then died within 28 days or even longer of any other cause – *any other cause*. Now subtract all those illusory 'Covid' deaths on the global data sheets from the 0.23 percent. What do you think you would be left with? *Zero*. A vaccination has never been successfully developed for a so-called coronavirus. They have all failed at the animal testing stage when they caused hypersensitivity to what they were claiming to protect against and made the impact of a disease far worse. Cult-owned vaccine corporations got around that problem this time by bypassing animal trials, going straight to humans and making the length of the 'trials' before the public rollout as short as they could get away with. Normally it takes five to ten years or more to develop vaccines that still cause demonstrable harm to many people and that's without including the long-term effects that are never officially connected to the vaccination. 'Covid' non-vaccines have been officially produced and approved in a matter of months from a standing start and part of the reason is that (a) they were developed before the 'Covid' hoax began and (b) they are based on computer programs and not natural sources. Official non-trials were so short that government agencies gave *emergency*, not full, approval. 'Trials' were not even completed and full approval cannot be secured until they are. Public 'Covid vaccination' is actually a *continuation of the trial*. Drug company 'trials' are not scheduled to end until 2023 by which time a lot of people are going to be dead. Data on which government agencies gave this emergency approval was supplied by the Big Pharma corporations themselves in the form of Pfizer/BioNTech, AstraZeneca, Moderna, Johnson & Johnson, and

others, and this is the case with all vaccines. By its very nature *emergency* approval means drug companies do not have to prove that the 'vaccine' is 'safe and effective'. How could they with trials way short of complete? Government regulators only have to *believe* that they *could* be safe and effective. It is criminal manipulation to get products in circulation with no testing worth the name. Agencies giving that approval are infested with Big Pharma-connected place-people and they act in the interests of Big Pharma (the Cult) and not the public about whom they do not give a damn.

More human lab rats

'Covid vaccines' produced in record time by Pfizer/BioNTech and Moderna employ a technique *never approved before for use on humans*. They are known as mRNA 'vaccines' and inject a synthetic version of 'viral' mRNA or 'messenger RNA'. The key is in the term 'messenger'. The body works, or doesn't, on the basis of information messaging. Communications are constantly passing between and within the genetic system and the brain. Change those messages and you change the state of the body and even its very nature and you can change psychology and behaviour by the way the brain processes information. I think you are going to see significant changes in personality and perception of many people who have had the 'Covid vaccine' synthetic potions. Insider Aldous Huxley predicted the following in 1961 and mRNA 'vaccines' can be included in the term 'pharmacological methods':

> There will be, in the next generation or so, a pharmacological method of making people love their servitude, and producing dictatorship without tears, so to speak, producing a kind of painless concentration camp for entire societies, so that people will in fact have their own liberties taken away from them, but rather enjoy it, because they will be distracted from any desire to rebel by propaganda or brainwashing, or brainwashing enhanced by pharmacological methods. And this seems to be the final revolution.

Apologists claim that mRNA synthetic 'vaccines' don't change the DNA genetic blueprint because RNA does not affect DNA only the other way round. This is so disingenuous. A process called 'reverse

transcription' can convert RNA into DNA and be integrated into DNA in the cell nucleus. This was highlighted in December, 2020, by scientists at Harvard and Massachusetts Institute of Technology (MIT). Geneticists report that more than 40 percent of mammalian genomes results from reverse transcription. On the most basic level if messaging changes then that sequence must lead to changes in DNA which is receiving and transmitting those communications. How can introducing synthetic material into cells not change the cells where DNA is located? The process is known as transfection which is defined as 'a technique to insert foreign nucleic acid (DNA or RNA) into a cell, typically with the intention of altering the properties of the cell'. Researchers at the Sloan Kettering Institute in New York found that changes in messenger RNA can deactivate tumour-suppressing proteins and thereby promote cancer. This is what happens when you mess with messaging. 'Covid vaccine' maker Moderna was founded in 2010 by Canadian stem cell biologist Derrick J. Rossi after his breakthrough discovery in the field of transforming and reprogramming stem cells. These are neutral cells that can be programmed to become any cell including sperm cells. Moderna was therefore founded on the principle of genetic manipulation and has never produced any vaccine or drug before its genetically-manipulating synthetic 'Covid' shite. Look at the name – Mode-RNA or Modify-RNA. Another important point is that the US Supreme Court has ruled that genetically-modified DNA, or complementary DNA (cDNA) synthesized in the laboratory from messenger RNA, can be patented and owned. These psychopaths are doing this to the human body.

Cells replicate synthetic mRNA in the 'Covid vaccines' and in theory the body is tricked into making antigens which trigger antibodies to target the 'virus spike proteins' which as Dr Tom Cowan said have *never been seen*. Cut the crap and these 'vaccines' deliver *self-replicating* synthetic material to the cells with the effect of changing human DNA. The more of them you have the more that process is compounded while synthetic material is all the time self-replicating. 'Vaccine'-maker Moderna describes mRNA as 'like

software for the cell' and so they are messing with the body's software. What happens when you change the software in a computer? Everything changes. For this reason the Cult is preparing a production line of mRNA 'Covid vaccines' and a long list of excuses to use them as with all the 'variants' of a 'virus' never shown to exist. The plan is further to transfer the mRNA technique to other vaccines mostly given to children and young people. The cumulative consequences will be a transformation of human DNA through a constant infusion of synthetic genetic material which will kill many and change the rest. Now consider that governments that have given emergency approval for a vaccine that's not a vaccine; never been approved for humans before; had no testing worth the name; and the makers have been given immunity from prosecution for any deaths or adverse effects suffered by the public. The UK government awarded *permanent legal indemnity* to itself and its employees for harm done when a patient is being treated for 'Covid-19' or 'suspected Covid-19'. That is quite a thought when these are possible 'side-effects' from the 'vaccine' (they are not 'side', they are effects) listed by the US Food and Drug Administration:

Guillain-Barre syndrome; acute disseminated encephalomyelitis; transverse myelitis; encephalitis; myelitis; encephalomyelitis; meningoencephalitis; meningitis; encephalopathy; convulsions; seizures; stroke; narcolepsy; cataplexy; anaphylaxis; acute myocardial infarction (heart attack); myocarditis; pericarditis; autoimmune disease; death; implications for pregnancy, and birth outcomes; other acute demyelinating diseases; non anaphylactic allergy reactions; thrombocytopenia ; disseminated intravascular coagulation; venous thromboembolism; arthritis; arthralgia; joint pain; Kawasaki disease; multisystem inflammatory syndrome in children; vaccine enhanced disease. The latter is the way the 'vaccine' has the potential to make diseases far worse than they would otherwise be.

UK doctor and freedom campaigner Vernon Coleman described the conditions in this list as 'all unpleasant, most of them very serious, and you can't get more serious than death'. The thought that anyone at all has had the 'vaccine' in these circumstances is testament to the potential that humanity has for clueless, unquestioning, stupidity and for many that programmed stupidity has already been terminal.

An insider speaks

Dr Michael Yeadon is a former Vice President, head of research and Chief Scientific Adviser at vaccine giant Pfizer. Yeadon worked on the inside of Big Pharma, but that did not stop him becoming a vocal critic of 'Covid vaccines' and their potential for multiple harms, including infertility in women. By the spring of 2021 he went much further and even used the no, no, term 'conspiracy'. When you begin to see what is going on it is impossible not to do so. Yeadon spoke out in an interview with freedom campaigner James Delingpole and I mentioned earlier how he said that no one had samples of 'the virus'. He explained that the mRNA technique originated in the anti-cancer field and ways to turn on and off certain genes which could be advantageous if you wanted to stop cancer growing out of control. 'That's the origin of them. They are a very unusual application, really.' Yeadon said that treating a cancer patient with an aggressive procedure might be understandable if the alternative was dying, but it was quite another thing to use the same technique as a public health measure. Most people involved wouldn't catch the infectious agent you were vaccinating against and if they did they probably wouldn't die:

> If you are really using it as a public health measure you really want to as close as you can get to zero sides-effects … I find it odd that they chose techniques that were really cutting their teeth in the field of oncology and I'm worried that in using gene-based vaccines that have to be injected in the body and spread around the body, get taken up into some cells, and the regulators haven't quite told us which cells they get taken up into … you are going to be generating a wide range of responses … with multiple steps each of which could go well or badly.

I doubt the Cult intends it to go well. Yeadon said that you can put any gene you like into the body through the 'vaccine'. 'You can certainly give them a gene that would do them some harm if you wanted.' I was intrigued when he said that when used in the cancer field the technique could turn genes on and off. I explore this process in *The Answer* and with different genes having different functions you could create mayhem – physically and psychologically – if you turned the wrong ones on and the right ones off. I read reports of an experiment by researchers at the University of Washington's school of computer science and engineering in which they encoded DNA to infect computers. The body is itself a biological computer and if human DNA can inflict damage on a computer why can't the computer via synthetic material mess with the human body? It can. The Washington research team said it was possible to insert malicious malware into 'physical DNA strands' and corrupt the computer system of a gene sequencing machine as it 'reads gene letters and stores them as binary digits 0 and 1'. They concluded that hackers could one day use blood or spit samples to access computer systems and obtain sensitive data from police forensics labs or infect genome files. It is at this level of digital interaction that synthetic 'vaccines' need to be seen to get the full picture and that will become very clear later on. Michael Yeadon said it made no sense to give the 'vaccine' to younger people who were in no danger from the 'virus'. What was the benefit? It was all downside with potential effects:

> The fact that my government in what I thought was a civilised, rational country, is raining [the 'vaccine'] on people in their 30s and 40s, even my children in their 20s, they're getting letters and phone calls, I know this is not right and any of you doctors who are vaccinating you know it's not right, too. They are not at risk. They are not at risk from the disease, so you are now hoping that the side-effects are so rare that you get away with it. You don't give new technology … that you don't understand to 100 percent of the population.

Blood clot problems with the AstraZeneca 'vaccine' have been affecting younger people to emphasise the downside risks with no benefit. AstraZeneca's version, produced with Oxford University, does not use mRNA, but still gets its toxic cocktail inside cells where

it targets DNA. The Johnson & Johnson 'vaccine' which uses a similar technique has also produced blood clot effects to such an extent that the United States paused its use at one point. They are all 'gene therapy' (cell modification) procedures and not 'vaccines'. The truth is that once the content of these injections enter cells we have no idea what the effect will be. People can speculate and some can give very educated opinions and that's good. In the end, though, only the makers know what their potions are designed to do and even they won't know every last consequence. Michael Yeadon was scathing about doctors doing what they knew to be wrong. 'Everyone's mute', he said. Doctors in the NHS must know this was not right, coming into work and injecting people. 'I don't know how they sleep at night. I know I couldn't do it. I know that if I were in that position I'd have to quit.' He said he knew enough about toxicology to know this was not a good risk-benefit. Yeadon had spoken to seven or eight university professors and all except two would not speak out publicly. Their universities had a policy that no one said anything that countered the government and its medical advisors. They were afraid of losing their government grants. This is how intimidation has been used to silence the truth at every level of the system. I say silence, but these people could still speak out if they made that choice. Yeadon called them 'moral cowards' – 'This is about your children and grandchildren's lives and you have just buggered off and left it.'

'Variant' nonsense

Some of his most powerful comments related to the alleged 'variants' being used to instil more fear, justify more lockdowns, and introduce more 'vaccines'. He said government claims about 'variants' were nonsense. He had checked the alleged variant 'codes' and they were 99.7 percent identical to the 'original'. This was the human identity difference equivalent to putting a baseball cap on and off or wearing it the other way round. A 0.3 percent difference would make it impossible for that 'variant' to escape immunity from the 'original'. This made no sense of having new 'vaccines' for

'variants'. He said there would have to be at least a *30 percent* difference for that to be justified and even then he believed the immune system would still recognise what it was. Gates-funded 'variant modeller' and 'vaccine'-pusher John Edmunds might care to comment. Yeadon said drug companies were making new versions of the 'vaccine' as a 'top up' for 'variants'. Worse than that, he said, the 'regulators' around the world like the MHRA in the UK had got together and agreed that because 'vaccines' for 'variants' were so similar to the first 'vaccines' *they did not have to do safety studies*. How transparently sinister that is. This is when Yeadon said: 'There is a conspiracy here.' There was no need for another vaccine for 'variants' and yet we were told that there was and the country had shut its borders because of them. 'They are going into hundreds of millions of arms without passing 'go' or any regulator. Why did they do that? Why did they pick this method of making the vaccine?'

The reason had to be something bigger than that it seemed and 'it's not protection against the virus'. It's was a far bigger project that meant politicians and advisers were willing to do things and not do things that knowingly resulted in avoidable deaths – 'that's already happened when you think about lockdown and deprivation of health care for a year.' He spoke of people prepared to do something that results in the avoidable death of their fellow human beings and it not bother them. This is the penny-drop I have been working to get across for more than 30 years – the level of pure evil we are dealing with. Yeadon said his friends and associates could not believe there could be that much evil, but he reminded them of Stalin, Pol Pot and Hitler and of what Stalin had said: 'One death is a tragedy. A million? A statistic.' He could not think of a benign explanation for why you need top-up vaccines 'which I'm sure you don't' and for the regulators 'to just get out of the way and wave them through'. Why would the regulators do that when they were still wrestling with the dangers of the 'parent' vaccine? He was clearly shocked by what he had seen since the 'Covid' hoax began and now he was thinking the previously unthinkable:

If you wanted to depopulate a significant proportion of the world and to do it in a way that doesn't involve destruction of the environment with nuclear weapons, poisoning everyone with anthrax or something like that, and you wanted plausible deniability while you had a multi-year infectious disease crisis, I actually don't think you could come up with a better plan of work than seems to be in front of me. I can't say that's what they are going to do, but I can't think of a benign explanation why they are doing it.

He said he never thought that they would get rid of 99 percent of humans, but now he wondered. 'If you wanted to that this would be a hell of a way to do it – it would be unstoppable folks.' Yeadon had concluded that those who submitted to the 'vaccine' would be allowed to have some kind of normal life (but for how long?) while screws were tightened to coerce and mandate the last few percent. 'I think they'll put the rest of them in a prison camp. I wish I was wrong, but I don't think I am.' Other points he made included: There were no coronavirus vaccines then suddenly they all come along at the same time; we have no idea of the long term affect with trials so short; coercing or forcing people to have medical procedures is against the Nuremberg Code instigated when the Nazis did just that; people should at least delay having the 'vaccine'; a quick Internet search confirms that masks don't reduce respiratory viral transmission and 'the government knows that'; they have smashed civil society and they know that, too; two dozen peer-reviewed studies show no connection between lockdown and reducing deaths; he knew from personal friends the elite were still flying around and going on holiday while the public were locked down; the elite were not having the 'vaccines'. He was also asked if 'vaccines' could be made to target difference races. He said he didn't know, but the document by the Project for the New American Century in September, 2000, said developing 'advanced forms of biological warfare that can target *specific genotypes* may transform biological warfare from the realm of terror to a politically useful tool.' Oh, they're evil all right. Of that we can be *absolutely* sure.

Another cull of old people

We have seen from the CDC definition that the mRNA 'Covid vaccine' is not a vaccine and nor are the others that *claim* to reduce 'severity of symptoms' in *some* people, but not protect from infection or transmission. What about all the lies about returning to 'normal' if people were 'vaccinated'? If they are not claimed to stop infection and transmission of the alleged 'virus', how does anything change? This was all lies to manipulate people to take the jabs and we are seeing that now with masks and distancing still required for the 'vaccinated'. How did they think that elderly people with fragile health and immune responses were going to be affected by infusing their cells with synthetic material and other toxic substances? They *knew* that in the short and long term it would be devastating and fatal as the culling of the old that began with the first lockdowns was continued with the 'vaccine'. Death rates in care homes soared immediately residents began to be 'vaccinated' – infused with synthetic material. Brave and committed whistleblower nurses put their careers at risk by exposing this truth while the rest kept their heads down and their mouths shut to put their careers before those they are supposed to care for. A long-time American Certified Nursing Assistant who gave his name as James posted a video in which he described emotionally what happened in his care home when vaccination began. He said that during 2020 very few residents were sick with 'Covid' and no one died during the entire year; but shortly after the Pfizer mRNA injections 14 people died within two weeks and many others were near death. 'They're dropping like flies', he said. Residents who walked on their own before the shot could no longer and they had lost their ability to conduct an intelligent conversation. The home's management said the sudden deaths were caused by a 'super-spreader' of 'Covid-19'. Then how come, James asked, that residents who refused to take the injections were not sick? It was a case of inject the elderly with mRNA synthetic potions and blame their illness and death that followed on the 'virus'. James described what was happening in care homes as 'the greatest crime of genocide this country has ever seen'. Remember the NHS staff nurse from earlier who used the same

word ‘genocide’ for what was happening with the ‘vaccines’ and that it was an ‘act of human annihilation’. A UK care home whistleblower told a similar story to James about the effect of the ‘vaccine’ in deaths and ‘outbreaks’ of illness dubbed ‘Covid’ after getting the jab. She told how her care home management and staff had zealously imposed government regulations and no one was allowed to even question the official narrative let alone speak out against it. She said the NHS was even worse. Again we see the results of reframing. A worker at a local care home where I live said they had not had a single case of ‘Covid’ there for almost a year and when the residents were ‘vaccinated’ they had 19 positive cases in two weeks with eight dying.

It’s not the ‘vaccine’ – honest

The obvious cause and effect was being ignored by the media and most of the public. Australia’s health minister Greg Hunt (a former head of strategy at the World Economic Forum) was admitted to hospital after he had the ‘vaccine’. He was suffering according to reports from the skin infection ‘cellulitis’ and it must have been a severe case to have warranted days in hospital. Immediately the authorities said this was nothing to do with the ‘vaccine’ when an effect of some vaccines is a ‘cellulitis-like reaction’. We had families of perfectly healthy old people who died after the ‘vaccine’ saying that if only they had been given the ‘vaccine’ earlier they would still be alive. As a numbskull rating that is off the chart. A father of four ‘died of Covid’ at aged 48 when he was taken ill two days after having the ‘vaccine’. The man, a health administrator, had been ‘shielding during the pandemic’ and had ‘not really left the house’ until he went for the ‘vaccine’. Having the ‘vaccine’ and then falling ill and dying does not seem to have qualified as a possible cause and effect and ‘Covid-19’ went on his death certificate. His family said they had no idea how he ‘caught the virus’. A family member said: ‘Tragically, it could be that going for a vaccination ultimately led to him catching Covid …The sad truth is that they are never going to know where it came from.’ The family warned people to remember

that the virus still existed and was 'very real'. So was their stupidity. Nurses and doctors who had the first round of the 'vaccine' were collapsing, dying and ending up in a hospital bed while they or their grieving relatives were saying they'd still have the 'vaccine' again despite what happened. I kid you not. You mean if your husband returned from the dead he'd have the same 'vaccine' again that killed him??

Doctors at the VCU Medical Center in Richmond, Virginia, said the Johnson & Johnson 'vaccine' was to blame for a man's skin peeling off. Patient Richard Terrell said: 'It all just happened so fast. My skin peeled off. It's still coming off on my hands now.' He said it was stinging, burning and itching and when he bent his arms and legs it was very painful with 'the skin swollen and rubbing against itself'. Pfizer/BioNTech and Moderna vaccines use mRNA to change the cell while the Johnson & Johnson version uses DNA in a process similar to AstraZeneca's technique. Johnson & Johnson and AstraZeneca have both had their 'vaccines' paused by many countries after causing serious blood problems. Terrell's doctor Fnu Nutan said he could have died if he hadn't got medical attention. It sounds terrible so what did Nutan and Terrell say about the 'vaccine' now? Oh, they still recommend that people have it. A nurse in a hospital bed 40 minutes after the vaccination and unable to swallow due to throat swelling was told by a doctor that he lost mobility in his arm for 36 hours following the vaccination. What did he say to the ailing nurse? 'Good for you for getting the vaccination.' We are dealing with a serious form of cognitive dissonance madness in both public and medical staff. There is a remarkable correlation between those having the 'vaccine' and trumpeting the fact and suffering bad happenings shortly afterwards. Witold Rogiewicz, a Polish doctor, made a video of his 'vaccination' and ridiculed those who were questioning its safety and the intentions of Bill Gates: 'Vaccinate yourself to protect yourself, your loved ones, friends and also patients. And to mention quickly I have info for anti-vaxxers and anti-Coviders if you want to contact Bill Gates you can do this through me.' He further ridiculed the dangers of 5G. Days later he

was dead, but naturally the vaccination wasn't mentioned in the verdict of 'heart attack'.

Lies, lies and more lies

So many members of the human race have slipped into extreme states of insanity and unfortunately they include reframed doctors and nursing staff. Having a 'vaccine' and dying within minutes or hours is not considered a valid connection while death from any cause within 28 days or longer of a positive test with a test not testing for the 'virus' means 'Covid-19' goes on the death certificate. How could that 'vaccine'-death connection not have been made except by calculated deceit? US figures in the initial rollout period to February 12th, 2020, revealed that a third of the deaths reported to the CDC after 'Covid vaccines' happened within 48 hours. Five men in the UK suffered an 'extremely rare' blood clot problem after having the AstraZeneca 'vaccine', but no causal link was established said the Gates-funded Medicines and Healthcare products Regulatory Agency (MHRA) which had given the 'vaccine' emergency approval to be used. Former Pfizer executive Dr Michael Yeadon explained in his interview how the procedures could cause blood coagulation and clots. People who should have been at no risk were dying from blood clots in the brain and he said he had heard from medical doctor friends that people were suffering from skin bleeding and massive headaches. The AstraZeneca 'shot' was stopped by some 20 countries over the blood clotting issue and still the corrupt MHRA, the European Medicines Agency (EMA) and the World Health Organization said that it should continue to be given even though the EMA admitted that it 'still cannot rule out definitively' a link between blood clotting and the 'vaccine'. Later Marco Cavaleri, head of EMA vaccine strategy, said there was indeed a clear link between the 'vaccine' and thrombosis, but they didn't know why. So much for the trials showing the 'vaccine' is safe. Blood clots were affecting younger people who would be under virtually no danger from 'Covid' even if it existed which makes it all the more stupid and sinister.

The British government responded to public alarm by wheeling out June Raine, the terrifyingly weak infant school headmistress sound-alike who heads the UK MHRA drug 'regulator'. The idea that she would stand up to Big Pharma and government pressure is laughable and she told us that all was well in the same way that she did when allowing untested, never-used-on-humans-before, genetically-manipulating 'vaccines' to be exposed to the public in the first place. Mass lying is the new normal of the 'Covid' era. The MHRA later said 30 cases of rare blood clots had by then been connected with the AstraZeneca 'vaccine' (that means a lot more in reality) while stressing that the benefits of the jab in preventing 'Covid-19' outweighed any risks. A more ridiculous and disingenuous statement with callous disregard for human health it is hard to contemplate. Immediately after the mendacious 'all-clears' two hospital workers in Denmark experienced blood clots and cerebral haemorrhaging following the AstraZeneca jab and one died. Top Norwegian health official Pål Andre Holme said the 'vaccine' was the only common factor: 'There is nothing in the patient history of these individuals that can give such a powerful immune response … I am confident that the antibodies that we have found are the cause, and I see no other explanation than it being the vaccine which triggers it.' Strokes, a clot or bleed in the brain, were clearly associated with the 'vaccine' from word of mouth and whistleblower reports. Similar consequences followed with all these 'vaccines' that we were told were so safe and as the numbers grew by the day it was clear we were witnessing human carnage.

Learning the hard way

A woman interviewed by UKColumn told how her husband suffered dramatic health effects after the vaccine when he'd been in good health all his life. He went from being a little unwell to losing all feeling in his legs and experiencing 'excruciating pain'. Misdiagnosis followed twice at Accident and Emergency (an 'allergy' and 'sciatica') before he was admitted to a neurology ward where doctors said his serious condition had been caused by the

'vaccine'. Another seven 'vaccinated' people were apparently being treated on the same ward for similar symptoms. The woman said he had the 'vaccine' because they believed media claims that it was safe. 'I didn't think the government would give out a vaccine that does this to somebody; I believed they would be bringing out a vaccination that would be safe.' What a tragic way to learn that lesson. Another woman posted that her husband was transporting stroke patients to hospital on almost every shift and when he asked them if they had been 'vaccinated' for 'Covid' they all replied 'yes'. One had a 'massive brain bleed' the day after his second dose. She said her husband reported the 'just been vaccinated' information every time to doctors in A and E only for them to ignore it, make no notes and appear annoyed that it was even mentioned. This particular report cannot be verified, but it expresses a common theme that confirms the monumental underreporting of 'vaccine' consequences. Interestingly as the 'vaccines' and their brain blood clot/stroke consequences began to emerge the UK National Health Service began a publicity campaign telling the public what to do in the event of a stroke. A Scottish NHS staff nurse who quit in disgust in March, 2021, said:

I have seen traumatic injuries from the vaccine, they're not getting reported to the yellow card [adverse reaction] scheme, they're treating the symptoms, not asking why, why it's happening. It's just treating the symptoms and when you speak about it you're dismissed like you're crazy, I'm not crazy, I'm not crazy because every other colleague I've spoken to is terrified to speak out, they've had enough.

Videos appeared on the Internet of people uncontrollably shaking after the 'vaccine' with no control over muscles, limbs and even their face. A Scottish mother broke out in a severe rash all over her body almost immediately after she was given the AstraZeneca 'vaccine'. The pictures were horrific. Leigh King, a 41-year-old hairdresser from Lanarkshire said: 'Never in my life was I prepared for what I was about to experience … My skin was so sore and constantly hot … I have never felt pain like this …' But don't you worry, the 'vaccine' is perfectly safe. Then there has been the effect on medical

staff who have been pressured to have the 'vaccine' by psychopathic 'health' authorities and government. A London hospital consultant who gave the name K. Polyakova wrote this to the *British Medical Journal* or *BMJ*:

> I am currently struggling with … the failure to report the reality of the morbidity caused by our current vaccination program within the health service and staff population. The levels of sickness after vaccination is unprecedented and staff are getting very sick and some with neurological symptoms which is having a huge impact on the health service function. Even the young and healthy are off for days, some for weeks, and some requiring medical treatment. Whole teams are being taken out as they went to get vaccinated together.
>
> Mandatory vaccination in this instance is stupid, unethical and irresponsible when it comes to protecting our staff and public health. We are in the voluntary phase of vaccination, and encouraging staff to take an unlicensed product that is impacting on their immediate health … it is clearly stated that these vaccine products do not offer immunity or stop transmission. In which case why are we doing it?

Not to protect health that's for sure. Medical workers are lauded by governments for agenda reasons when they couldn't give a toss about them any more than they can for the population in general. Schools across America faced the same situation as they closed due to the high number of teachers and other staff with bad reactions to the Pfizer/BioNTech, Moderna, and Johnson & Johnson 'Covid vaccines' all of which were linked to death and serious adverse effects. The *BMJ* took down the consultant's comments pretty quickly on the grounds that they were being used to spread 'disinformation'. They were exposing the truth about the 'vaccine' was the real reason. The cover-up is breathtaking.

Hiding the evidence

The scale of the 'vaccine' death cover-up worldwide can be confirmed by comparing official figures with the personal experience of the public. I heard of many people in my community who died immediately or soon after the vaccine that would never appear in the media or even likely on the official totals of 'vaccine' fatalities and adverse reactions when only about ten percent are estimated to be

reported and I have seen some estimates as low as one percent in a Harvard study. In the UK alone by April 29th, 2021, some 757,654 adverse reactions had been officially reported from the Pfizer/BioNTech, Oxford/AstraZeneca and Moderna 'vaccines' with more than a thousand deaths linked to jabs and that means an estimated ten times this number in reality from a ten percent reporting rate percentage. That's seven million adverse reactions and 10,000 potential deaths and a one percent reporting rate would be ten times *those* figures. In 1976 the US government pulled the swine flu vaccine after *53* deaths. The UK data included a combined 10,000 eye disorders from the 'Covid vaccines' with more than 750 suffering visual impairment or blindness and again multiply by the estimated reporting percentages. As 'Covid cases' officially fell hospitals virtually empty during the 'Covid crisis' began to fill up with a range of other problems in the wake of the 'vaccine' rollout. The numbers across America have also been catastrophic. Deaths linked to *all* types of vaccine increased by *6,000 percent* in the first quarter of 2021 compared with 2020. A 39-year-old woman from Ogden, Utah, died four days after receiving a second dose of Moderna's 'Covid vaccine' when her liver, heart and kidneys all failed despite the fact that she had no known medical issues or conditions. Her family sought an autopsy, but Dr Erik Christensen, Utah's chief medical examiner, said proving vaccine injury as a cause of death almost never happened. He could think of only one instance where an autopsy would name a vaccine as the official cause of death and that would be anaphylaxis where someone received a vaccine and died almost instantaneously. 'Short of that, it would be difficult for us to definitively say this is the vaccine,' Christensen said. If that is true this must be added to the estimated ten percent (or far less) reporting rate of vaccine deaths and serious reactions and the conclusion can only be that vaccine deaths and serious reactions – including these 'Covid' potions' – are phenomenally understated in official figures. The same story can be found everywhere. Endless accounts of deaths and serious reactions among the public, medical

and care home staff while official figures did not even begin to reflect this.

Professional script-reader Dr David Williams, a 'top public-health official' in Ontario, Canada, insulted our intelligence by claiming only four serious adverse reactions and no deaths from the more than 380,000 vaccine doses then given. This bore no resemblance to what people knew had happened in their owns circles and we had Dirk Huyer in charge of getting millions vaccinated in Ontario while at the same time he was Chief Coroner for the province investigating causes of death including possible death from the vaccine. An aide said he had stepped back from investigating deaths, but evidence indicated otherwise. Rosemary Frei, who secured a Master of Science degree in molecular biology at the Faculty of Medicine at Canada's University of Calgary before turning to investigative journalism, was one who could see that official figures for 'vaccine' deaths and reactions made no sense. She said that doctors seldom reported adverse events and when people got really sick or died after getting a vaccination they would attribute that to anything except the vaccines. It had been that way for years and anyone who wondered aloud whether the 'Covid vaccines' or other shots cause harm is immediately branded as 'anti-vax' and 'anti-science'. This was 'career-threatening' for health professionals. Then there was the huge pressure to support the push to 'vaccinate' billions in the quickest time possible. Frei said:

> So that's where we're at today. More than half a million vaccine doses have been given to people in Ontario alone. The rush is on to vaccinate all 15 million of us in the province by September. And the mainstream media are screaming for this to be sped up even more. That all adds up to only a very slim likelihood that we're going to be told the truth by officials about how many people are getting sick or dying from the vaccines.

What is true of Ontario is true of everywhere.

They KNEW – and still did it

The authorities knew what was going to happen with multiple deaths and adverse reactions. The UK government's Gates-funded

and Big Pharma-dominated Medicines and Healthcare products Regulatory Agency (MHRA) hired a company to employ AI in compiling the projected reactions to the 'vaccine' that would otherwise be uncountable. The request for applications said: 'The MHRA urgently seeks an Artificial Intelligence (AI) software tool to process the expected high volume of Covid-19 vaccine Adverse Drug Reaction …' This was from the agency, headed by the disingenuous June Raine, that gave the 'vaccines' emergency approval and the company was hired before the first shot was given. 'We are going to kill and maim you – is that okay?' 'Oh, yes, perfectly fine – I'm very grateful, thank you, doctor.' The range of 'Covid vaccine' adverse reactions goes on for page after page in the MHRA criminally underreported 'Yellow Card' system and includes affects to eyes, ears, skin, digestion, blood and so on. Raine's MHRA amazingly claimed that the 'overall safety experience … is so far as expected from the clinical trials'. The death, serious adverse effects, deafness and blindness were *expected?* When did they ever mention that? If these human tragedies were expected then those that gave approval for the use of these 'vaccines' must be guilty of crimes against humanity including murder – a definition of which is 'killing a person with malice aforethought or with recklessness manifesting extreme indifference to the value of human life.' People involved at the MHRA, the CDC in America and their equivalent around the world must go before Nuremberg trials to answer for their callous inhumanity. We are only talking here about the immediate effects of the 'vaccine'. The longer-term impact of the DNA synthetic manipulation is the main reason they are so hysterically desperate to inoculate the entire global population in the shortest possible time.

Africa and the developing world are a major focus for the 'vaccine' depopulation agenda and a mass vaccination sales-pitch is underway thanks to caring people like the Rockefellers and other Cult assets. The Rockefeller Foundation, which pre-empted the 'Covid pandemic' in a document published in 2010 that 'predicted' what happened a decade later, announced an initial $34.95 million grant in February, 2021, 'to ensure more equitable access to Covid-19

testing and vaccines' among other things in Africa in collaboration with '24 organizations, businesses, and government agencies'. The pan-Africa initiative would focus on 10 countries: Burkina Faso, Ethiopia, Ghana, Kenya, Nigeria, Rwanda, South Africa, Tanzania, Uganda, and Zambia'. Rajiv Shah, President of the Rockefeller Foundation and former administrator of CIA-controlled USAID, said that if Africa was not mass-vaccinated (to change the DNA of its people) it was a 'threat to all of humanity' and not fair on Africans. When someone from the Rockefeller Foundation says they want to do something to help poor and deprived people and countries it is time for a belly-laugh. They are doing this out of the goodness of their 'heart' because 'vaccinating' the entire global population is what the 'Covid' hoax set out to achieve. Official 'decolonisation' of Africa by the Cult was merely a prelude to financial colonisation on the road to a return to physical colonisation. The 'vaccine' is vital to that and the sudden and convenient death of the 'Covid' sceptic president of Tanzania can be seen in its true light. A lot of people in Africa are aware that this is another form of colonisation and exploitation and they need to stand their ground.

The 'vaccine is working' scam

A potential problem for the Cult was that the 'vaccine' is meant to change human DNA and body messaging and not to protect anyone from a 'virus' never shown to exist. The vaccine couldn't work because it was not designed to work and how could they make it *appear* to be working so that more people would have it? This was overcome by lowering the amplification rate of the PCR test to produce fewer 'cases' and therefore fewer 'deaths'. Some of us had been pointing out since March, 2020, that the amplification rate of the test not testing for the 'virus' had been made artificially high to generate positive tests which they could call 'cases' to justify lockdowns. The World Health Organization recommended an absurdly high 45 amplification cycles to ensure the high positives required by the Cult and then remained silent on the issue until January 20th, 2021 – Biden's Inauguration Day. This was when the

'vaccinations' were seriously underway and on that day the WHO recommended after discussions with America's CDC that laboratories *lowered their testing amplification*. Dr David Samadi, a certified urologist and health writer, said the WHO was encouraging all labs to reduce their cycle count for PCR tests. He said the current cycle was much too high and was 'resulting in any particle being declared a positive case'. Even one mainstream news report I saw said this meant the number of 'Covid' infections may have been 'dramatically inflated'. Oh, just a little bit. The CDC in America issued new guidance to laboratories in April, 2021, to use 28 cycles *but only for 'vaccinated' people*. The timing of the CDC/WHO interventions were cynically designed to make it appear the 'vaccines' were responsible for falling cases and deaths when the real reason can be seen in the following examples. New York's state lab, the Wadsworth Center, identified 872 positive tests in July, 2020, based on a threshold of 40 cycles. When the figure was lowered to 35 cycles *43 percent* of the 872 were no longer 'positives'. At 30 cycles the figure was 63 percent. A Massachusetts lab found that between *85 to 90 percent* of people who tested positive in July with a cycle threshold of 40 would be negative at 30 cycles, Ashish Jha, MD, director of the Harvard Global Health Institute, said: 'I'm really shocked that it could be that high … Boy, does it really change the way we need to be thinking about testing.' I'm shocked that I could see the obvious in the spring of 2020, with no medical background, and most medical professionals still haven't worked it out. No, that's not shocking – it's terrifying.

Three weeks after the WHO directive to lower PCR cycles the London *Daily Mail* ran this headline: 'Why ARE Covid cases plummeting? New infections have fallen 45% in the US and 30% globally in the past 3 weeks but experts say vaccine is NOT the main driver because only 8% of Americans and 13% of people worldwide have received their first dose.' They acknowledged that the drop could not be attributed to the 'vaccine', but soon this morphed throughout the media into the 'vaccine' has caused cases and deaths to fall when it was the PCR threshold. In December, 2020, there was

chaos at English Channel ports with truck drivers needing negative 'Covid' tests before they could board a ferry home for Christmas. The government wanted to remove the backlog as fast as possible and they brought in troops to do the 'testing'. Out of 1,600 drivers just *36* tested positive and the rest were given the all clear to cross the Channel. I guess the authorities thought that 36 was the least they could get away with without the unquestioning catching on. The amplification trick which most people believed in the absence of information in the mainstream applied more pressure on those refusing the 'vaccine' to succumb when it 'obviously worked'. The truth was the exact opposite with deaths in care homes soaring with the 'vaccine' and in Israel the term used was 'skyrocket'. A re-analysis of published data from the Israeli Health Ministry led by Dr Hervé Seligmann at the Medicine Emerging Infectious and Tropical Diseases at Aix-Marseille University found that Pfizer's 'Covid vaccine' killed 'about 40 times more [elderly] people than the disease itself would have killed' during a five-week vaccination period and *260 times* more younger people than would have died from the 'virus' even according to the manipulated 'virus' figures. Dr Seligmann and his co-study author, Haim Yativ, declared after reviewing the Israeli 'vaccine' death data: 'This is a new Holocaust.'

Then, in mid-April, 2021, after vast numbers of people worldwide had been 'vaccinated', the story changed with clear coordination. The UK government began to prepare the ground for more future lockdowns when Nuremberg-destined Boris Johnson told yet another whopper. He said that cases had fallen because of *lockdowns* not 'vaccines'. Lockdowns are irrelevant when *there is no 'virus'* and the test and fraudulent death certificates are deciding the number of 'cases' and 'deaths'. Study after study has shown that lockdowns don't work and instead kill and psychologically destroy people. Meanwhile in the United States Anthony Fauci and Rochelle Walensky, the ultra-Zionist head of the CDC, peddled the same line. More lockdown was the answer and not the 'vaccine', a line repeated on cue by the moron that is Canadian Prime Minister Justin Trudeau. Why all the hysteria to get everyone 'vaccinated' if lockdowns and

not 'vaccines' made the difference? None of it makes sense on the face of it. Oh, but it does. The Cult wants lockdowns *and* the 'vaccine' and if the 'vaccine' is allowed to be seen as the total answer lockdowns would no longer be justified when there are still livelihoods to destroy. 'Variants' and renewed upward manipulation of PCR amplification are planned to instigate never-ending lockdown *and* more 'vaccines'.

You *must* have it – we're desperate

Israel, where the Jewish and Arab population are ruled by the Sabbatian Cult, was the front-runner in imposing the DNA-manipulating 'vaccine' on its people to such an extent that Jewish refusers began to liken what was happening to the early years of Nazi Germany. This would seem to be a fantastic claim. Why would a government of Jewish people be acting like the Nazis did? If you realise that the Sabbatian Cult was behind the Nazis and that Sabbatians hate Jews the pieces start to fit and the question of why a 'Jewish' government would treat Jews with such callous disregard for their lives and freedom finds an answer. Those controlling the government of Israel *aren't Jewish* – they're Sabbatian. Israeli lawyer Tamir Turgal was one who made the Nazi comparison in comments to German lawyer Reiner Fuellmich who is leading a class action lawsuit against the psychopaths for crimes against humanity. Turgal described how the Israeli government was vaccinating children and pregnant women on the basis that there was no evidence that this was dangerous when they had no evidence that it *wasn't* dangerous either. They just had no evidence. This was medical experimentation and Turgal said this breached the Nuremberg Code about medical experimentation and procedures requiring informed consent and choice. Think about that. A Nuremberg Code developed because of Nazi experimentation on Jews and others in concentration camps by people like the evil-beyond-belief Josef Mengele is being breached by the *Israeli* government; but when you know that it's a *Sabbatian* government along with its intelligence and military agencies like Mossad, Shin Bet and the Israeli Defense Forces, and that Sabbatians

were the force behind the Nazis, the kaleidoscope comes into focus. What have we come to when Israeli Jews are suing their government for violating the Nuremberg Code by essentially making Israelis subject to a medical experiment using the controversial 'vaccines'? It's a shocker that this has to be done in the light of what happened in Nazi Germany. The Anshe Ha-Emet, or 'People of the Truth', made up of Israeli doctors, lawyers, campaigners and public, have launched a lawsuit with the International Criminal Court. It says:

> When the heads of the Ministry of Health as well as the prime minister presented the vaccine in Israel and began the vaccination of Israeli residents, the vaccinated were not advised, that, in practice, they are taking part in a medical experiment and that their consent is required for this under the Nuremberg Code.

The irony is unbelievable, but easily explained in one word: Sabbatians. The foundation of Israeli 'Covid' apartheid is the 'green pass' or 'green passport' which allows Jews and Arabs who have had the DNA-manipulating 'vaccine' to go about their lives – to work, fly, travel in general, go to shopping malls, bars, restaurants, hotels, concerts, gyms, swimming pools, theatres and sports venues, while non-'vaccinated' are banned from all those places and activities. Israelis have likened the 'green pass' to the yellow stars that Jews in Nazi Germany were forced to wear – the same as the yellow stickers that a branch of UK supermarket chain Morrisons told exempt mask-wears they had to display when shopping. How very sensitive. The Israeli system is blatant South African-style apartheid on the basis of compliance or non-compliance to fascism rather than colour of the skin. How appropriate that the Sabbatian Israeli government was so close to the pre-Mandela apartheid regime in Pretoria. The Sabbatian-instigated 'vaccine passport' in Israel is planned for everywhere. Sabbatians struck a deal with Pfizer that allowed them to lead the way in the percentage of a national population infused with synthetic material and the result was catastrophic. Israeli freedom activist Shai Dannon told me how chairs were appearing on beaches that said 'vaccinated only'. Health Minister Yuli Edelstein said that anyone unwilling or unable to get

the jabs that 'confer immunity' will be 'left behind'. The man's a liar. Not even the makers claim the 'vaccines' confer immunity. When you see those figures of 'vaccine' deaths these psychopaths were saying that you must take the chance the 'vaccine' will kill you or maim you while knowing it will change your DNA or lockdown for you will be permanent. That's fascism. The Israeli parliament passed a law to allow personal information of the non-vaccinated to be shared with local and national authorities for three months. This was claimed by its supporters to be a way to 'encourage' people to be vaccinated. Hadas Ziv from Physicians for Human Rights described this as a 'draconian law which crushed medical ethics and the patient rights'. But that's the idea, the Sabbatians would reply.

Your papers, please

Sabbatian Israel was leading what has been planned all along to be a global 'vaccine pass' called a 'green passport' without which you would remain in permanent lockdown restriction and unable to do anything. This is how badly – *desperately* – the Cult is to get everyone 'vaccinated'. The term and colour 'green' was not by chance and related to the psychology of fusing the perception of the green climate hoax with the 'Covid' hoax and how the 'solution' to both is the same Great Reset. Lying politicians, health officials and psychologists denied there were any plans for mandatory vaccinations or restrictions based on vaccinations, but they knew that was exactly what was meant to happen with governments of all countries reaching agreements to enforce a global system. 'Free' Denmark and 'free' Sweden unveiled digital vaccine certification. Cyprus, Czech Republic, Estonia, Greece, Hungary, Iceland, Italy, Poland, Portugal, Slovakia, and Spain have all committed to a vaccine passport system and the rest including the whole of the EU would follow. The satanic UK government will certainly go this way despite mendacious denials and at the time of writing it is trying to manipulate the public into having the 'vaccine' so they could go abroad on a summer holiday. How would that work without something to prove you had the synthetic toxicity injected into you?

Documents show that the EU's European Commission was moving towards 'vaccine certificates' in 2018 and 2019 before the 'Covid' hoax began. They knew what was coming. Abracadabra – Ursula von der Leyen, the German President of the Commission, announced in March, 2021, an EU 'Digital Green Certificate' – green again – to track the public's 'Covid status'. The passport sting is worldwide and the Far East followed the same pattern with South Korea ruling that only those with 'vaccination' passports – again the *green* pass – would be able to 'return to their daily lives'.

Bill Gates has been preparing for this 'passport' with other Cult operatives for years and beyond the paper version is a Gates-funded 'digital tattoo' to identify who has been vaccinated and who hasn't. The 'tattoo' is reported to include a substance which is externally readable to confirm who has been vaccinated. This is a bio-luminous light-generating enzyme (think fireflies) called … *Luciferase*. Yes, named after the Cult 'god' Lucifer the 'light bringer' of whom more to come. Gates said he funded the readable tattoo to ensure children in the developing world were vaccinated and no one was missed out. He cares so much about poor kids as we know. This was just the cover story to develop a vaccine tagging system for everyone on the planet. Gates has been funding the ID2020 'alliance' to do just that in league with other lovely people at Microsoft, GAVI, the Rockefeller Foundation, Accenture and IDEO.org. He said in interviews in March, 2020, before any 'vaccine' publicly existed, that the world must have a globalised digital certificate to track the 'virus' and who had been vaccinated. Gates knew from the start that the mRNA vaccines were coming and when they would come and that the plan was to tag the 'vaccinated' to marginalise the intelligent and stop them doing anything including travel. Evil just doesn't suffice. Gates was exposed for offering a $10 million bribe to the Nigerian House of Representatives to invoke compulsory 'Covid' vaccination of all Nigerians. Sara Cunial, a member of the Italian Parliament, called Gates a 'vaccine criminal'. She urged the Italian President to hand him over to the International Criminal Court for crimes against

humanity and condemned his plans to 'chip the human race' through ID2020.

You know it's a long-planned agenda when war criminal and Cult gofer Tony Blair is on the case. With the scale of arrogance only someone as dark as Blair can muster he said: 'Vaccination in the end is going to be your route to liberty.' Blair is a disgusting piece of work and he confirms that again. The media has given a lot of coverage to a bloke called Charlie Mullins, founder of London's biggest independent plumbing company, Pimlico Plumbers, who has said he won't employ anyone who has not been vaccinated or have them go to any home where people are not vaccinated. He said that if he had his way no one would be allowed to walk the streets if they have not been vaccinated. Gates was cheering at the time while I was alerting the white coats. The plan is that people will qualify for 'passports' for having the first two doses and then to keep it they will have to have all the follow ups and new ones for invented 'variants' until human genetics is transformed and many are dead who can't adjust to the changes. Hollywood celebrities – the usual propaganda stunt – are promoting something called the WELL Health-Safety Rating to verify that a building or space has 'taken the necessary steps to prioritize the health and safety of their staff, visitors and other stakeholders'. They included Lady Gaga, Jennifer Lopez, Michael B. Jordan, Robert DeNiro, Venus Williams, Wolfgang Puck, Deepak Chopra and 17th Surgeon General Richard Carmona. Yawn. WELL Health-Safety has big connections with China. Parent company Delos is headed by former Goldman Sachs partner Paul Scialla. This is another example – and we will see so many others – of using the excuse of 'health' to dictate the lives and activities of the population. I guess one confirmation of the 'safety' of buildings is that only 'vaccinated' people can go in, right?

Electronic concentration camps

I wrote decades ago about the plans to restrict travel and here we are for those who refuse to bow to tyranny. This can be achieved in one go with air travel if the aviation industry makes a blanket decree.

The ‘vaccine’ and guaranteed income are designed to be part of a global version of China’s social credit system which tracks behaviour 24/7 and awards or deletes ‘credits’ based on whether your behaviour is supported by the state or not. I mean your entire lifestyle – what you do, eat, say, everything. Once your credit score falls below a certain level consequences kick in. In China tens of millions have been denied travel by air and train because of this. All the locations and activities denied to refusers by the ‘vaccine’ passports will be included in one big mass ban on doing almost anything for those that don’t bow their head to government. It’s beyond fascist and a new term is required to describe its extremes – I guess fascist technocracy will have to do. The way the Chinese system of technological – technocratic – control is sweeping the West can be seen in the Los Angeles school system and is planned to be expanded worldwide. Every child is required to have a ‘Covid’-tracking app scanned daily before they can enter the classroom. The so-called Daily Pass tracking system is produced by Gates’ Microsoft which I’m sure will shock you rigid. The pass will be scanned using a barcode (one step from an inside-the-body barcode) and the information will include health checks, ‘Covid’ tests and vaccinations. Entry codes are for one specific building only and access will only be allowed if a student or teacher has a negative test with a test not testing for the ‘virus’, has no symptoms of anything alleged to be related to ‘Covid’ (symptoms from a range of other illness), and has a temperature under 100 degrees. No barcode, no entry, is planned to be the case for everywhere and not only schools.

Kids are being psychologically prepared to accept this as ‘normal’ their whole life which is why what they can impose in schools is so important to the Cult and its gofers. Long-time American freedom campaigner John Whitehead of the Rutherford Institute was not exaggerating when he said: ‘Databit by databit, we are building our own electronic concentration camps.’ Canada under its Cult gofer prime minister Justin Trudeau has taken a major step towards the real thing with people interned against their will if they test positive with a test not testing for the ‘virus’ when they arrive at a Canadian

airport. They are jailed in internment hotels often without food or water for long periods and with many doors failing to lock there have been sexual assaults. The interned are being charged sometimes $2,000 for the privilege of being abused in this way. Trudeau is fully on board with the Cult and says the 'Covid pandemic' has provided an opportunity for a global 'reset' to permanently change Western civilisation. His number two, Deputy Prime Minister Chrystia Freeland, is a trustee of the World Economic Forum and a Rhodes Scholar. The Trudeau family have long been servants of the Cult. See *The Biggest Secret* and Cathy O'Brien's book *Trance-Formation of America* for the horrific background to Trudeau's father Pierre Trudeau another Canadian prime minister. Hide your fascism behind the façade of a heart-on-the-sleeve liberal. It's a well-honed Cult technique.

What can the 'vaccine' *really* do?

We have a 'virus' never shown to exist and 'variants' of the 'virus' that have also never been shown to exist except, like the 'original', as computer-generated fictions. Even if you believe there's a 'virus' the 'case' to 'death' rate is in the region of 0.23 to 0.15 percent and those 'deaths' are concentrated among the very old around the same average age that people die anyway. In response to this lack of threat (in truth none) psychopaths and idiots, knowingly and unknowingly answering to Gates and the Cult, are seeking to 'vaccinate' every man, woman and child on Planet Earth. Clearly the 'vaccine' is not about 'Covid' – none of this ever has been. So what is it all about *really*? Why the desperation to infuse genetically-manipulating synthetic material into everyone through mRNA fraudulent 'vaccines' with the intent of doing this over and over with the excuses of 'variants' and other 'virus' inventions? Dr Sherri Tenpenny, an osteopathic medical doctor in the United States, has made herself an expert on vaccines and their effects as a vehement campaigner against their use. Tenpenny was board certified in emergency medicine, the director of a level two trauma centre for 12 years, and moved to Cleveland in 1996 to start an integrative

medicine practice which has treated patients from all 50 states and some 17 other countries. Weaning people off pharmaceutical drugs is a speciality.

She became interested in the consequences of vaccines after attending a meeting at the National Vaccine Information Center in Washington DC in 2000 where she 'sat through four days of listening to medical doctors and scientists and lawyers and parents of vaccine injured kids' and asked: 'What's going on?' She had never been vaccinated and never got ill while her father was given a list of vaccines to be in the military and was 'sick his entire life'. The experience added to her questions and she began to examine vaccine documents from the Centers for Disease Control (CDC). After reading the first one, the 1998 version of *The General Recommendations of Vaccination*, she thought: 'This is it?' The document was poorly written and bad science and Tenpenny began 20 years of research into vaccines that continues to this day. She began her research into 'Covid vaccines' in March, 2020, and she describes them as 'deadly'. For many, as we have seen, they already have been. Tenpenny said that in the first 30 days of the 'vaccine' rollout in the United States there had been more than 40,000 adverse events reported to the vaccine adverse event database. A document had been delivered to her the day before that was 172 pages long. 'We have over 40,000 adverse events; we have over 3,100 cases of [potentially deadly] anaphylactic shock; we have over 5,000 neurological reactions.' Effects ranged from headaches to numbness, dizziness and vertigo, to losing feeling in hands or feet and paraesthesia which is when limbs 'fall asleep' and people have the sensation of insects crawling underneath their skin. All this happened in the first 30 days and remember that only about *ten percent* (or far less) of adverse reactions and vaccine-related deaths are estimated to be officially reported. Tenpenny said:

> So can you think of one single product in any industry, any industry, for as long as products have been made on the planet that within 30 days we have 40,000 people complaining of side effects that not only is still on the market but … we've got paid actors telling us how great

they are for getting their vaccine. We're offering people $500 if they will just get their vaccine and we've got nurses and doctors going; 'I got the vaccine, I got the vaccine'.

Tenpenny said they were not going to be 'happy dancing folks' when they began to suffer Bell's palsy (facial paralysis), neuropathies, cardiac arrhythmias and autoimmune reactions that kill through a blood disorder. 'They're not going to be so happy, happy then, but we're never going to see pictures of those people' she said. Tenpenny described the 'vaccine' as 'a well-designed killing tool'.

No off-switch

Bad as the initial consequences had been Tenpenny said it would be maybe 14 months before we began to see the 'full ravage' of what is going to happen to the 'Covid vaccinated' with full-out consequences taking anything between two years and 20 years to show. You can understand why when you consider that variations of the 'Covid vaccine' use mRNA (messenger RNA) to in theory activate the immune system to produce protective antibodies without using the actual 'virus'. How can they when it's a computer program and they've never isolated what they claim is the 'real thing'? Instead they use *synthetic* mRNA. They are inoculating synthetic material into the body which through a technique known as the Trojan horse is absorbed into cells to change the nature of DNA. Human DNA is changed by an infusion of messenger RNA and with each new 'vaccine' of this type it is changed even more. Say so and you are banned by Cult Internet platforms. The contempt the contemptuous Mark Zuckerberg has for the truth and human health can be seen in an internal Facebook video leaked to the Project Veritas investigative team in which he said of the 'Covid vaccines': '… I share some caution on this because we just don't know the long term side-effects of basically modifying people's DNA and RNA.' At the same time this disgusting man's Facebook was censoring and banning anyone saying exactly the same. He must go before a Nuremberg trial for crimes against humanity when he *knows* that he

is censoring legitimate concerns and denying the right of informed consent on behalf of the Cult that owns him. People have been killed and damaged by the very 'vaccination' technique he cast doubt on himself when they may not have had the 'vaccine' with access to information that he denied them. The plan is to have at least annual 'Covid vaccinations', add others to deal with invented 'variants', and change all other vaccines into the mRNA system. Pfizer executives told shareholders at a virtual Barclays Global Healthcare Conference in March, 2021, that the public may need a third dose of 'Covid vaccine', plus regular yearly boosters and the company planned to hike prices to milk the profits in a 'significant opportunity for our vaccine'. These are the professional liars, cheats and opportunists who are telling you their 'vaccine' is safe. Given this volume of mRNA planned to be infused into the human body and its ability to then replicate we will have a transformation of human genetics from biological to synthetic biological – exactly the long-time Cult plan for reasons we'll see – and many will die. Sherri Tenpenny said of this replication:

> It's like having an on-button but no off-button and that whole mechanism … they actually give it a name and they call it the Trojan horse mechanism, because it allows that [synthetic] virus and that piece of that [synthetic] virus to get inside of your cells, start to replicate and even get inserted into other parts of your DNA as a Trojan-horse.

Ask the overwhelming majority of people who have the 'vaccine' what they know about the contents and what they do and they would reply: 'The government says it will stop me getting the virus.' Governments give that false impression on purpose to increase take-up. You can read Sherri Tenpenny's detailed analysis of the health consequences in her blog at Vaxxter.com, but in summary these are some of them. She highlights the statement by Bill Gates about how human beings can become their own 'vaccine manufacturing machine'. The man is insane. ['Vaccine'-generated] 'antibodies' carry synthetic messenger RNA into the cells and the damage starts, Tenpenny contends, and she says that lungs can be adversely affected through varying degrees of pus and bleeding which

obviously affects breathing and would be dubbed 'Covid-19'. Even more sinister was the impact of 'antibodies' on macrophages, a white blood cell of the immune system. They consist of Type 1 and Type 2 which have very different functions. She said Type 1 are 'hyper-vigilant' white blood cells which 'gobble up' bacteria etc. However, in doing so, this could cause inflammation and in extreme circumstances be fatal. She says these affects are mitigated by Type 2 macrophages which kick in to calm down the system and stop it going rogue. They clear up dead tissue debris and reduce inflammation that the Type 1 'fire crews' have caused. Type 1 kills the infection and Type 2 heals the damage, she says. This is her punchline with regard to 'Covid vaccinations': She says that mRNA 'antibodies' block Type 2 macrophages by attaching to them and deactivating them. This meant that when the Type 1 response was triggered by infection there was nothing to stop that getting out of hand by calming everything down. There's an on-switch, but no off-switch, she says. What follows can be 'over and out, see you when I see you'.

Genetic suicide

Tenpenny also highlights the potential for autoimmune disease – the body attacking itself – which has been associated with vaccines since they first appeared. Infusing a synthetic foreign substance into cells could cause the immune system to react in a panic believing that the body is being overwhelmed by an invader (it is) and the consequences can again be fatal. There is an autoimmune response known as a 'cytokine storm' which I have likened to a homeowner panicked by an intruder and picking up a gun to shoot randomly in all directions before turning the fire on himself. The immune system unleashes a storm of inflammatory response called cytokines to a threat and the body commits hara-kiri. The lesson is that you mess with the body's immune response at your peril and these 'vaccines' seriously – fundamentally – mess with immune response. Tenpenny refers to a consequence called anaphylactic shock which is a severe and highly dangerous allergic reaction when the immune system

floods the body with chemicals. She gives the example of having a bee sting which primes the immune system and makes it sensitive to those chemicals. When people are stung again maybe years later the immune response can be so powerful that it leads to anaphylactic shock. Tenpenny relates this 'shock' with regard to the 'Covid vaccine' to something called polyethylene glycol or PEG. Enormous numbers of people have become sensitive to this over decades of use in a whole range of products and processes including food, drink, skin creams and 'medicine'. Studies have claimed that some 72 percent of people have antibodies triggered by PEG compared with two percent in the 1960s and allergic hypersensitive reactions to this become a gathering cause for concern. Tenpenny points out that the 'mRNA vaccine' is coated in a 'bubble' of polyethylene glycol which has the potential to cause anaphylactic shock through immune sensitivity. Many reports have appeared of people reacting this way after having the 'Covid vaccine'. What do we think is going to happen as humanity has more and more of these 'vaccines'? Tenpenny said: 'All these pictures we have seen with people with these rashes … these weepy rashes, big reactions on their arms and things like that – it's an acute allergic reaction most likely to the polyethylene glycol that you've been previously primed and sensitised to.'

Those who have not studied the conspiracy and its perpetrators at length might think that making the population sensitive to PEG and then putting it in these 'vaccines' is just a coincidence. It is not. It is instead testament to how carefully and coldly-planned current events have been and the scale of the conspiracy we are dealing with. Tenpenny further explains that the 'vaccine' mRNA procedure can breach the blood-brain barrier which protects the brain from toxins and other crap that will cause malfunction. In this case they could make two proteins corrupt brain function to cause Amyotrophic lateral sclerosis (ALS) , a progressive nervous system disease leading to loss of muscle control, and frontal lobe degeneration – Alzheimer's and dementia. Immunologist J. Bart Classon published a paper connecting mRNA 'vaccines' to prion

disease which can lead to Alzheimer's and other forms of neurogenerative disease while others have pointed out the potential to affect the placenta in ways that make women infertile. This will become highly significant in the next chapter when I will discuss other aspects of this non-vaccine that relate to its nanotechnology and transmission from the injected to the uninjected.

Qualified in idiocy

Tenpenny describes how research has confirmed that these 'vaccine'-generated antibodies can interact with a range of other tissues in the body and attack many other organs including the lungs. 'This means that if you have a hundred people standing in front of you that all got this shot they could have a hundred different symptoms.' Anyone really think that Cult gofers like the Queen, Tony Blair, Christopher Whitty, Anthony Fauci, and all the other psychopaths have really had this 'vaccine' in the pictures we've seen? Not a bloody chance. Why don't doctors all tell us about all these dangers and consequences of the 'Covid vaccine'? Why instead do they encourage and pressure patients to have the shot? Don't let's think for a moment that doctors and medical staff can't be stupid, lazy, and psychopathic and that's without the financial incentives to give the jab. Tenpenny again:

> Some people are going to die from the vaccine directly but a large number of people are going to start to get horribly sick and get all kinds of autoimmune diseases 42 days to maybe a year out. What are they going to do, these stupid doctors who say; 'Good for you for getting that vaccine.' What are they going to say; 'Oh, it must be a mutant, we need to give an extra dose of that vaccine.'
>
> Because now the vaccine, instead of one dose or two doses we need three or four because the stupid physicians aren't taking the time to learn anything about it. If I can learn this sitting in my living room reading a 19 page paper and several others so can they. There's nothing special about me, I just take the time to do it.

Remember how Sara Kayat, the NHS and TV doctor, said that the 'Covid vaccine' would '100 percent prevent hospitalisation and death'. Doctors can be idiots like every other profession and they

should not be worshipped as infallible. They are not and far from it. Behind many medical and scientific 'experts' lies an uninformed prat trying to hide themselves from you although in the 'Covid' era many have failed to do so as with UK narrative-repeating 'TV doctor' Hilary Jones. Pushing back against the minority of proper doctors and scientists speaking out against the 'vaccine' has been the entire edifice of the Cult global state in the form of governments, medical systems, corporations, mainstream media, Silicon Valley, and an army of compliant doctors, medical staff and scientists willing to say anything for money and to enhance their careers by promoting the party line. If you do that you are an 'expert' and if you won't you are an 'anti-vaxxer' and 'Covidiot'. The pressure to be 'vaccinated' is incessant. We have even had reports claiming that the 'vaccine' can help cure cancer and Alzheimer's and make the lame walk. I am waiting for the announcement that it can bring you coffee in the morning and cook your tea. Just as the symptoms of 'Covid' seem to increase by the week so have the miracles of the 'vaccine'. American supermarket giant Kroger Co. offered nearly 500,000 employees in 35 states a $100 bonus for having the 'vaccine' while donut chain Krispy Kreme promised 'vaccinated' customers a free glazed donut every day for the rest of 2021. Have your DNA changed and you will get a doughnut although we might not have to give you them for long. Such offers and incentives confirm the desperation.

Perhaps the worse vaccine-stunt of them all was UK 'Health' Secretary Matt-the-prat Hancock on live TV after watching a clip of someone being 'vaccinated' when the roll-out began. Hancock faked tears so badly it was embarrassing. Brain-of-Britain Piers Morgan, the lockdown-supporting, 'vaccine' supporting, 'vaccine' passport-supporting, TV host played along with Hancock – 'You're quite emotional about that' he said in response to acting so atrocious it would have been called out at a school nativity which will presumably today include Mary and Jesus in masks, wise men keeping their camels six feet apart, and shepherds under tent arrest. System-serving Morgan tweeted this: 'Love the idea of covid vaccine passports for everywhere: flights, restaurants, clubs, football, gyms,

shops etc. It's time covid-denying, anti-vaxxer loonies had their bullsh*t bluff called & bar themselves from going anywhere that responsible citizens go.' If only I could aspire to his genius. To think that Morgan, who specialises in shouting over anyone he disagrees with, was lauded as a free speech hero when he lost his job after storming off the set of his live show like a child throwing his dolly out of the pram. If he is a free speech hero we are in real trouble. I have no idea what 'bullsh*t' means, by the way, the * throws me completely.

The Cult is desperate to infuse its synthetic DNA-changing concoction into everyone and has been using every lie, trick and intimidation to do so. The question of '*Why*?' we shall now address.

CHAPTER TEN

Human 2.0

I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted –
Alan Turing (1912-1954), the 'Father of artificial intelligence'

I have been exposing for decades the plan to transform the human body from a biological to a synthetic-biological state. The new human that I will call Human 2.0 is planned to be connected to artificial intelligence and a global AI 'Smart Grid' that would operate as one global system in which AI would control everything from your fridge to your heating system to your car to your mind. Humans would no longer be 'human', but post-human and sub-human, with their thinking and emotional processes replaced by AI.

What I said sounded crazy and beyond science fiction and I could understand that. To any balanced, rational, mind it *is* crazy. Today, however, that world is becoming reality and it puts the 'Covid vaccine' into its true context. Ray Kurzweil is the ultra-Zionist 'computer scientist, inventor and futurist' and co-founder of the Singularity University. Singularity refers to the merging of humans with machines or 'transhumanism'. Kurzweil has said humanity would be connected to the cyber 'cloud' in the period of the ever-recurring year of 2030:

Our thinking … will be a hybrid of biological and non-biological thinking … humans will be able to extend their limitations and 'think in the cloud' … We're going to put gateways to the

cloud in our brains ... We're going to gradually merge and enhance ourselves ... In my view, that's the nature of being human – we transcend our limitations. As the technology becomes vastly superior to what we are then the small proportion that is still human gets smaller and smaller and smaller until it's just utterly negligible.

They are trying to sell this end-of-humanity-as-we-know-it as the next stage of 'evolution' when we become super-human and 'like the gods'. They are lying to you. Shocked, eh? The population, and again especially the young, have been manipulated into addiction to technologies designed to enslave them for life. First they induced an addiction to smartphones (holdables); next they moved to technology on the body (wearables); and then began the invasion of the body (implantables). I warned way back about the plan for microchipped people and we are now entering that era. We should not be diverted into thinking that this refers only to chips we can see. Most important are the nanochips known as smart dust, neural dust and nanobots which are far too small to be seen by the human eye. Nanotechnology is everywhere, increasingly in food products, and released into the atmosphere by the geoengineering of the skies funded by Bill Gates to 'shut out the Sun' and 'save the planet from global warming'. Gates has been funding a project to spray millions of tonnes of chalk (calcium carbonate) into the stratosphere over Sweden to 'dim the Sun' and cool the Earth. Scientists warned the move could be disastrous for weather systems in ways no one can predict and opposition led to the Swedish space agency announcing that the 'experiment' would not be happening as planned in the summer of 2021; but it shows where the Cult is going with dimming the impact of the Sun and there's an associated plan to change the planet's atmosphere. Who gives psychopath Gates the right to dictate to the entire human race and dismantle planetary systems? The world will not be safe while this man is at large.

The global warming hoax has made the Sun, like the gas of life, something to fear when both are essential to good health and human survival (more inversion). The body transforms sunlight into vital vitamin D through a process involving … *cholesterol*. This is the cholesterol we are also told to fear. We are urged to take Big Pharma

statin drugs to reduce cholesterol and it's all systematic. Reducing cholesterol means reducing vitamin D uptake with all the multiple health problems that will cause. At least if you take statins long term it saves the government from having to pay you a pension. The delivery system to block sunlight is widely referred to as chemtrails although these have a much deeper agenda, too. They appear at first to be contrails or condensation trails streaming from aircraft into cold air at high altitudes. Contrails disperse very quickly while chemtrails do not and spread out across the sky before eventually their content falls to earth. Many times I have watched aircraft cross-cross a clear blue sky releasing chemtrails until it looks like a cloudy day. Chemtrails contain many things harmful to humans and the natural world including toxic heavy metals, aluminium (see Alzheimer's) and nanotechnology. Ray Kurzweil reveals the reason without actually saying so: 'Nanobots will infuse all the matter around us with information. Rocks, trees, everything will become these intelligent creatures.' How do you deliver that? *From the sky*. Self-replicating nanobots would connect everything to the Smart Grid. The phenomenon of Morgellons disease began in the chemtrail era and the correlation has led to it being dubbed the 'chemtrail disease'. Self-replicating fibres appear in the body that can be pulled out through the skin. Morgellons fibres continue to grow outside the body and have a form of artificial intelligence. I cover this at greater length in *Phantom Self*.

'Vaccine' operating system

'Covid vaccines' with their self-replicating synthetic material are also designed to make the connection between humanity and Kurzweil's 'cloud'. American doctor and dedicated campaigner for truth, Carrie Madej, an Internal Medicine Specialist in Georgia with more than 20 years medical experience, has highlighted the nanotechnology aspect of the fake 'vaccines'. She explains how one of the components in at least the Moderna and Pfizer synthetic potions are 'lipid nanoparticles' which are 'like little tiny computer bits' – a 'sci-fi substance' known as nanobots and hydrogel which can be 'triggered

at any moment to deliver its payload' and act as 'biosensors'. The synthetic substance had 'the ability to accumulate data from your body like your breathing, your respiration, thoughts and emotions, all kind of things' and each syringe could carry a *million* nanobots:

This substance because it's like little bits of computers in your body, crazy, but it's true, it can do that, [and] obviously has the ability to act through Wi-Fi. It can receive and transmit energy, messages, frequencies or impulses. That issue has never been addressed by these companies. What does that do to the human?

Just imagine getting this substance in you and it can react to things all around you, the 5G, your smart device, your phones, what is happening with that? What if something is triggering it, too, like an impulse, a frequency? We have something completely foreign in the human body.

Madej said her research revealed that electromagnetic (EMF) frequencies emitted by phones and other devices had increased dramatically in the same period of the 'vaccine' rollout and she was seeing more people with radiation problems as 5G and other electromagnetic technology was expanded and introduced to schools and hospitals. She said she was 'floored with the EMF coming off' the devices she checked. All this makes total sense and syncs with my own work of decades when you think that Moderna refers in documents to its mRNA 'vaccine' as an 'operating system':

Recognizing the broad potential of mRNA science, we set out to create an mRNA technology platform that functions very much like an operating system on a computer. It is designed so that it can plug and play interchangeably with different programs. In our case, the 'program' or 'app' is our mRNA drug – the unique mRNA sequence that codes for a protein …

… Our MRNA Medicines – 'The 'Software Of Life': When we have a concept for a new mRNA medicine and begin research, fundamental components are already in place. Generally, the only thing that changes from one potential mRNA medicine to another is the coding region – the actual genetic code that instructs ribosomes to make protein. Utilizing these instruction sets gives our investigational mRNA medicines a software-like quality. We also have the ability to combine different mRNA sequences encoding for different proteins in a single mRNA investigational medicine.

Who needs a real 'virus' when you can create a computer version to justify infusing your operating system into the entire human race on the road to making living, breathing people into cyborgs? What is missed with the 'vaccines' is the *digital* connection between synthetic material and the body that I highlighted earlier with the study that hacked a computer with human DNA. On one level the body is digital, based on mathematical codes, and I'll have more about that in the next chapter. Those who ridiculously claim that mRNA 'vaccines' are not designed to change human genetics should explain the words of Dr Tal Zaks, chief medical officer at Moderna, in a 2017 TED talk. He said that over the last 30 years 'we've been living this phenomenal digital scientific revolution, and I'm here today to tell you, that we are actually *hacking the software of life*, and that it's changing the way we think about prevention and treatment of disease':

In every cell there's this thing called messenger RNA, or mRNA for short, that transmits the critical information from the DNA in our genes to the protein, which is really the stuff we're all made out of. This is the critical information that determines what the cell will do. So we think about it as an operating system. So if you could change that, if you could introduce a line of code, or change a line of code, it turns out, that has profound implications for everything, from the flu to cancer.

Zaks should more accurately have said that this has profound implications for the human genetic code and the nature of DNA. Communications within the body go both ways and not only one. But, hey, no, the 'Covid vaccine' will not affect your genetics. Cult fact-checkers say so even though the man who helped to develop the mRNA technique says that it does. Zaks said in 2017:

If you think about what it is we're trying to do. We've taken information and our understanding of that information and how that information is transmitted in a cell, and we've taken our understanding of medicine and how to make drugs, and we're fusing the two. We think of it as information therapy.

I have been writing for decades that the body is an information field communicating with itself and the wider world. This is why

radiation which is information can change the information field of body and mind through phenomena like 5G and change their nature and function. 'Information therapy' means to change the body's information field and change the way it operates. DNA is a receiver-transmitter of information and can be mutated by information like mRNA synthetic messaging. Technology to do this has been ready and waiting in the underground bases and other secret projects to be rolled out when the 'Covid' hoax was played. 'Trials' of such short and irrelevant duration were only for public consumption. When they say the 'vaccine' is 'experimental' that is not true. It may appear to be 'experimental' to those who don't know what's going on, but the trials have already been done to ensure the Cult gets the result it desires. Zaks said that it took decades to sequence the human genome, completed in 2003, but now they could do it in a week. By 'they' he means scientists operating in the public domain. In the secret projects they were sequencing the genome in a week long before even 2003.

Deluge of mRNA

Highly significantly the Moderna document says the guiding premise is that if using mRNA as a medicine works for one disease then it should work for many diseases. They were leveraging the flexibility afforded by their platform and the fundamental role mRNA plays in protein synthesis to pursue mRNA medicines for a broad spectrum of diseases. Moderna is confirming what I was saying through 2020 that multiple 'vaccines' were planned for 'Covid' (and later invented 'variants') and that previous vaccines would be converted to the mRNA system to infuse the body with massive amounts of genetically-manipulating synthetic material to secure a transformation to a synthetic-biological state. The 'vaccines' are designed to kill stunning numbers as part of the long-exposed Cult depopulation agenda and transform the rest. Given this is the goal you can appreciate why there is such hysterical demand for every human to be 'vaccinated' for an alleged 'disease' that has an estimated 'infection' to 'death' ratio of 0.23-0.15 percent. As I write

children are being given the 'vaccine' in trials (their parents are a disgrace) and ever-younger people are being offered the vaccine for a 'virus' that even if you believe it exists has virtually zero chance of harming them. Horrific effects of the 'trials' on a 12-year-old girl were revealed by a family member to be serious brain and gastric problems that included a bowel obstruction and the inability to swallow liquids or solids. She was unable to eat or drink without throwing up, had extreme pain in her back, neck and abdomen, and was paralysed from the waist down which stopped her urinating unaided. When the girl was first taken to hospital doctors said it was all in her mind. She was signed up for the 'trial' by her parents for whom no words suffice. None of this 'Covid vaccine' insanity makes any sense unless you see what the 'vaccine' really is – a body-changer. Synthetic biology or 'SynBio' is a fast-emerging and expanding scientific discipline which includes everything from genetic and molecular engineering to electrical and computer engineering. Synthetic biology is defined in these ways:

- A multidisciplinary area of research that seeks to create new biological parts, devices, and systems, or to redesign systems that are already found in nature.

- The use of a mixture of physical engineering and genetic engineering to create new (and therefore synthetic) life forms.

- An emerging field of research that aims to combine the knowledge and methods of biology, engineering and related disciplines in the design of chemically-synthesized DNA to create organisms with novel or enhanced characteristics and traits (synthetic organisms including humans).

We now have synthetic blood, skin, organs and limbs being developed along with synthetic body parts produced by 3D printers. These are all elements of the synthetic human programme and this comment by Kurzweil's co-founder of the Singularity University,

Peter Diamandis, can be seen in a whole new light with the 'Covid' hoax and the sanctions against those that refuse the 'vaccine':

Anybody who is going to be resisting the progress forward [to transhumanism] is going to be resisting evolution and, fundamentally, they will die out. It's not a matter of whether it's good or bad. It's going to happen.

'Resisting evolution'? What absolute bollocks. The arrogance of these people is without limit. His 'it's going to happen' mantra is another way of saying 'resistance is futile' to break the spirit of those pushing back and we must not fall for it. Getting this genetically-transforming 'vaccine' into everyone is crucial to the Cult plan for total control and the desperation to achieve that is clear for anyone to see. Vaccine passports are a major factor in this and they, too, are a form of resistance is futile. It's NOT. The paper funded by the Rockefeller Foundation for the 2013 'health conference' in China said:

We will interact more with artificial intelligence. The use of robotics, bio-engineering to augment human functioning is already well underway and will advance. Re-engineering of humans into potentially separate and unequal forms through genetic engineering or mixed human-robots raises debates on ethics and equality.

A new demography is projected to emerge after 2030 [that year again] of technologies (robotics, genetic engineering, nanotechnology) producing robots, engineered organisms, 'nanobots' and artificial intelligence (AI) that can self-replicate. Debates will grow on the implications of an impending reality of human designed life.

What is happening today is so long planned. The world army enforcing the will of the world government is intended to be a robot army, not a human one. Today's military and its technologically 'enhanced' troops, pilotless planes and driverless vehicles are just stepping stones to that end. Human soldiers are used as Cult fodder and its time they woke up to that and worked for the freedom of the population instead of their own destruction and their family's destruction – the same with the police. Join us and let's sort this out. The phenomenon of enforce my own destruction is widespread in the 'Covid' era with Woker 'luvvies' in the acting and entertainment

industries supporting 'Covid' rules which have destroyed their profession and the same with those among the public who put signs on the doors of their businesses 'closed due to Covid – stay safe' when many will never reopen. It's a form of masochism and most certainly insanity.

Transgender = transhumanism

When something explodes out of nowhere and is suddenly everywhere it is always the Cult agenda and so it is with the tidal wave of claims and demands that have infiltrated every aspect of society under the heading of 'transgenderism'. The term 'trans' is so 'in' and this is the dictionary definition:

A prefix meaning 'across', 'through', occurring … in loanwords from Latin, used in particular for denoting movement or conveyance from place to place (transfer; transmit; transplant) or complete change (transform; transmute), or to form adjectives meaning 'crossing', 'on the other side of', or 'going beyond' the place named (transmontane; transnational; trans-Siberian).

Transgender means to go beyond gender and transhuman means to go beyond human. Both are aspects of the Cult plan to transform the human body to a synthetic state with *no gender*. Human 2.0 is not designed to procreate and would be produced technologically with no need for parents. The new human would mean the end of parents and so men, and increasingly women, are being targeted for the deletion of their rights and status. Parental rights are disappearing at an ever-quickening speed for the same reason. The new human would have no need for men or women when there is no procreation and no gender. Perhaps the transgender movement that appears to be in a permanent state of frenzy might now contemplate on how it is being used. This was never about transgender rights which are only the interim excuse for confusing gender, particularly in the young, on the road to *fusing* gender. Transgender activism is not an end; it is a *means* to an end. We see again the technique of creative destruction in which you destroy the status quo to 'build back better' in the form that you want. The gender status quo had to be

destroyed by persuading the Cult-created Woke mentality to believe that you can have 100 genders or more. A programme for 9 to 12 year olds produced by the Cult-owned BBC promoted the 100 genders narrative. The very idea may be the most monumental nonsense, but it is not what is true that counts, only what you can make people *believe* is true. Once the gender of 2 + 2 = 4 has been dismantled through indoctrination, intimidation and 2 + 2 = 5 then the new no-gender normal can take its place with Human 2.0. Aldous Huxley revealed the plan in his prophetic *Brave New World* in 1932:

> Natural reproduction has been done away with and children are created, decanted', and raised in 'hatcheries and conditioning centres'. From birth, people are genetically designed to fit into one of five castes, which are further split into 'Plus' and 'Minus' members and designed to fulfil predetermined positions within the social and economic strata of the World State.

How could Huxley know this in 1932? For the same reason George Orwell knew about the Big Brother state in 1948, Cult insiders I have quoted knew about it in 1969, and I have known about it since the early 1990s. If you are connected to the Cult or you work your balls off to uncover the plan you can predict the future. The process is simple. If there is a plan for the world and nothing intervenes to stop it then it will happen. Thus if you communicate the plan ahead of time you are perceived to have predicted the future, but you haven't. You have revealed the plan which without intervention will become the human future. The whole reason I have done what I have is to alert enough people to inspire an intervention and maybe at last that time has come with the Cult and its intentions now so obvious to anyone with a brain in working order.

The future is here

Technological wombs that Huxley described to replace parent procreation are already being developed and they are only the projects we know about in the public arena. Israeli scientists told *The Times of Israel* in March, 2021, that they have grown 250-cell embryos

into mouse foetuses with fully formed organs using artificial wombs in a development they say could pave the way for gestating humans outside the womb. Professor Jacob Hanna of the Weizmann Institute of Science said:

We took mouse embryos from the mother at day five of development, when they are just of 250 cells, and had them in the incubator from day five until day 11, by which point they had grown all their organs.

By day 11 they make their own blood and have a beating heart, a fully developed brain. Anybody would look at them and say, 'this is clearly a mouse foetus with all the characteristics of a mouse.' It's gone from being a ball of cells to being an advanced foetus.

A special liquid is used to nourish embryo cells in a laboratory dish and they float on the liquid to duplicate the first stage of embryonic development. The incubator creates all the right conditions for its development, Hanna said. The liquid gives the embryo 'all the nutrients, hormones and sugars they need' along with a custom-made electronic incubator which controls gas concentration, pressure and temperature. The cutting-edge in the underground bases and other secret locations will be light years ahead of that, however, and this was reported by the London *Guardian* in 2017:

We are approaching a biotechnological breakthrough. Ectogenesis, the invention of a complete external womb, could completely change the nature of human reproduction. In April this year, researchers at the Children's Hospital of Philadelphia announced their development of an artificial womb.

The article was headed 'Artificial wombs could soon be a reality. What will this mean for women?' What would it mean for children is an even bigger question. No mother to bond with only a machine in preparation for a life of soulless interaction and control in a world governed by machines (see the *Matrix* movies). Now observe the calculated manipulations of the 'Covid' hoax as human interaction and warmth has been curtailed by distancing, isolation and fear with people communicating via machines on a scale never seen before.

These are all dots in the same picture as are all the personal assistants, gadgets and children's toys through which kids and adults communicate with AI as if it is human. The AI 'voice' on Sat-Nav should be included. All these things are psychological preparation for the Cult endgame. Before you can make a physical connection with AI you have to make a psychological connection and that is what people are being conditioned to do with this ever gathering human-AI interaction. Movies and TV programmes depicting the transhuman, robot dystopia relate to a phenomenon known as 'pre-emptive programming' in which the world that is planned is portrayed everywhere in movies, TV and advertising. This is conditioning the conscious and subconscious mind to become familiar with the planned reality to dilute resistance when it happens for real. What would have been a shock such is the change is made less so. We have young children put on the road to transgender transition surgery with puberty blocking drugs at an age when they could never be able to make those life-changing decisions.

Rachel Levine, a professor of paediatrics and psychiatry who believes in treating children this way, became America's highest-ranked openly-transgender official when she was confirmed as US Assistant Secretary at the Department of Health and Human Services after being nominated by Joe Biden (the Cult). Activists and governments press for laws to deny parents a say in their children's transition process so the kids can be isolated and manipulated into agreeing to irreversible medical procedures. A Canadian father Robert Hoogland was denied bail by the Vancouver Supreme Court in 2021 and remained in jail for breaching a court order that he stay silent over his young teenage daughter, a minor, who was being offered life-changing hormone therapy without parental consent. At the age of 12 the girl's 'school counsellor' said she may be transgender, referred her to a doctor and told the school to treat her like a boy. This is another example of state-serving schools imposing ever more control over children's lives while parents have ever less.

Contemptible and extreme child abuse is happening all over the world as the Cult gender-fusion operation goes into warp-speed.

Why the war on men – and now women?

The question about what artificial wombs mean for women should rightly be asked. The answer can be seen in the deletion of women's rights involving sport, changing rooms, toilets and status in favour of people in male bodies claiming to identify as women. I can identify as a mountain climber, but it doesn't mean I can climb a mountain any more than a biological man can be a biological woman. To believe so is a triumph of belief over factual reality which is the very perceptual basis of everything Woke. Women's sport is being destroyed by allowing those with male bodies who say they identify as female to 'compete' with girls and women. Male body 'women' dominate 'women's' competition with their greater muscle mass, bone density, strength and speed. With that disadvantage sport for women loses all meaning. To put this in perspective nearly 300 American high school boys can run faster than the quickest woman sprinter in the world. Women are seeing their previously protected spaces invaded by male bodies simply because they claim to identify as women. That's all they need to do to access all women's spaces and activities under the Biden 'Equality Act' that destroys equality for women with the usual Orwellian Woke inversion. Male sex offenders have already committed rapes in women's prisons after claiming to identify as women to get them transferred. Does this not matter to the Woke 'equality' hypocrites? Not in the least. What matters to Cult manipulators and funders behind transgender activists is to advance gender fusion on the way to the no-gender 'human'. When you are seeking to impose transparent nonsense like this, or the 'Covid' hoax, the only way the nonsense can prevail is through censorship and intimidation of dissenters, deletion of factual information, and programming of the unquestioning, bewildered and naive. You don't have to scan the world for long to see that all these things are happening.

Many women's rights organisations have realised that rights and status which took such a long time to secure are being eroded and that it is systematic. Kara Dansky of the global Women's Human Rights Campaign said that Biden's transgender executive order immediately he took office, subsequent orders, and Equality Act legislation that followed 'seek to erase women and girls in the law as a category'. *Exactly*. I said during the long ago-started war on men (in which many women play a crucial part) that this was going to turn into a war on them. The Cult is phasing out *both* male and female genders. To get away with that they are brought into conflict so they are busy fighting each other while the Cult completes the job with no unity of response. Unity, people, *unity*. We need unity everywhere. Transgender is the only show in town as the big step towards the no-gender human. It's not about rights for transgender people and never has been. Woke political correctness is deleting words relating to genders to the same end. Wokers believe this is to be 'inclusive' when the opposite is true. They are deleting words describing gender because gender *itself* is being deleted by Human 2.0. Terms like 'man', 'woman', 'mother' and 'father' are being deleted in the universities and other institutions to be replaced by the *no*-gender, not trans-gender, 'individuals' and 'guardians'. Women's rights campaigner Maria Keffler of Partners for Ethical Care said: 'Children are being taught from kindergarten upward that some boys have a vagina, some girls have a penis, and that kids can be any gender they want to be.' Do we really believe that suddenly countries all over the world at the same time had the idea of having drag queens go into schools or read transgender stories to very young children in the local library? It's coldly-calculated confusion of gender on the way to the fusion of gender. Suzanne Vierling, a psychologist from Southern California, made another important point:

Yesterday's slave woman who endured gynecological medical experiments is today's girl-child being butchered in a booming gender-transitioning sector. Ovaries removed, pushing her into menopause and osteoporosis, uncharted territory, and parents' rights and authority decimated.

The erosion of parental rights is a common theme in line with the Cult plans to erase the very concept of parents and 'ovaries removed, pushing her into menopause' means what? Those born female lose the ability to have children – another way to discontinue humanity as we know it.

Eliminating Human 1.0 (before our very eyes)

To pave the way for Human 2.0 you must phase out Human 1.0. This is happening through plummeting sperm counts and making women infertile through an onslaught of chemicals, radiation (including smartphones in pockets of men) and mRNA 'vaccines'. Common agriculture pesticides are also having a devastating impact on human fertility. I have been tracking collapsing sperm counts in the books for a long time and in 2021 came a book by fertility scientist and reproductive epidemiologist Shanna Swan, *Count Down: How Our Modern World Is Threatening Sperm Counts, Altering Male and Female Reproductive Development and Imperiling the Future of the Human Race*. She reports how the global fertility rate dropped by *half* between 1960 and 2016 with America's birth rate 16 percent below where it needs to be to sustain the population. Women are experiencing declining egg quality, more miscarriages, and more couples suffer from infertility. Other findings were an increase in erectile dysfunction, infant boys developing more genital abnormalities, male problems with conception, and plunging levels of the male hormone testosterone which would explain why so many men have lost their backbone and masculinity. This has been very evident during the 'Covid' hoax when women have been prominent among the Pushbackers and big strapping blokes have bowed their heads, covered their faces with a nappy and quietly submitted. Mind control expert Cathy O'Brien also points to how global education introduced the concept of 'we're all winners' in sport and classrooms: 'Competition was defused, and it in turn defused a sense of fighting back.' This is another version of the 'equity' doctrine in which you drive down rather than raise up. What a contrast in Cult-controlled China with its global ambitions

where the government published plans in January, 2021, to 'cultivate masculinity' in boys from kindergarten through to high school in the face of a 'masculinity crisis'. A government adviser said boys would be soon become 'delicate, timid and effeminate' unless action was taken. Don't expect any similar policy in the targeted West. A 2006 study showed that a 65-year-old man in 2002 had testosterone levels *15 percent* lower than a 65-year-old man in 1987 while a 2020 study found a similar story with young adults and adolescents. Men are getting prescriptions for testosterone replacement therapy which causes an even greater drop in sperm count with up to 99 percent seeing sperm counts drop to zero during the treatment. More sperm is defective and malfunctioning with some having two heads or not pursuing an egg.

A class of *synthetic* chemicals known as phthalates are being blamed for the decline. These are found everywhere in plastics, shampoos, cosmetics, furniture, flame retardants, personal care products, pesticides, canned foods and even receipts. Why till receipts? Everyone touches them. Let no one delude themselves that all this is not systematic to advance the long-time agenda for human body transformation. Phthalates mimic hormones and disrupt the hormone balance causing testosterone to fall and genital birth defects in male infants. Animals and fish have been affected in the same way due to phthalates and other toxins in rivers. When fish turn gay or change sex through chemicals in rivers and streams it is a pointer to why there has been such an increase in gay people and the sexually confused. It doesn't matter to me what sexuality people choose to be, but if it's being affected by chemical pollution and consumption then we need to know. Does anyone really think that this is not connected to the transgender agenda, the war on men and the condemnation of male 'toxic masculinity'? You watch this being followed by 'toxic femininity'. It's already happening. When breastfeeding becomes 'chest-feeding', pregnant women become pregnant people along with all the other Woke claptrap you know that the world is going insane and there's a Cult scam in progress. Transgender activists are promoting the Cult agenda while Cult

billionaires support and fund the insanity as they laugh themselves to sleep at the sheer stupidity for which humans must be infamous in galaxies far, far away.

'Covid vaccines' and female infertility

We can now see why the 'vaccine' has been connected to potential infertility in women. Dr Michael Yeadon, former Vice President and Chief Scientific Advisor at Pfizer, and Dr Wolfgang Wodarg in Germany, filed a petition with the European Medicines Agency in December, 2020, urging them to stop trials for the Pfizer/BioNTech shot and all other mRNA trials until further studies had been done. They were particularly concerned about possible effects on fertility with 'vaccine'-produced antibodies attacking the protein Syncytin-1 which is responsible for developing the placenta. The result would be infertility 'of indefinite duration' in women who have the 'vaccine' with the placenta failing to form. Section 10.4.2 of the Pfizer/BioNTech trial protocol says that pregnant women or those who might become so should not have mRNA shots. Section 10.4 warns men taking mRNA shots to 'be abstinent from heterosexual intercourse' and not to donate sperm. The UK government said that it *did not know* if the mRNA procedure had an effect on fertility. *Did not know?* These people have to go to jail. UK government advice did not recommend at the start that pregnant women had the shot and said they should avoid pregnancy for at least two months after 'vaccination'. The 'advice' was later updated to pregnant women should only have the 'vaccine' if the benefits outweighed the risks to mother and foetus. What the hell is that supposed to mean? Then 'spontaneous abortions' began to appear and rapidly increase on the adverse reaction reporting schemes which include only a fraction of adverse reactions. Thousands and ever-growing numbers of 'vaccinated' women are describing changes to their menstrual cycle with heavier blood flow, irregular periods and menstruating again after going through the menopause – all links to reproduction effects. Women are passing blood clots and the lining of their uterus while men report erectile dysfunction and blood effects. Most

significantly of all *un*vaccinated women began to report similar menstrual changes after interaction with *'vaccinated'* people and men and children were also affected with bleeding noses, blood clots and other conditions. 'Shedding' is when vaccinated people can emit the content of a vaccine to affect the unvaccinated, but this is different. 'Vaccinated' people were not shedding a 'live virus' allegedly in 'vaccines' as before because the fake 'Covid vaccines' involve synthetic material and other toxicity. Doctors exposing what is happening prefer the term 'transmission' to shedding. Somehow those that have had the shots are transmitting effects to those that haven't. Dr Carrie Madej said the nano-content of the 'vaccines' can 'act like an antenna' to others around them which fits perfectly with my own conclusions. This 'vaccine' transmission phenomenon was becoming known as the book went into production and I deal with this further in the Postscript.

Vaccine effects on sterility are well known. The World Health Organization was accused in 2014 of sterilising millions of women in Kenya with the evidence confirmed by the content of the vaccines involved. The same WHO behind the 'Covid' hoax admitted its involvement for more than ten years with the vaccine programme. Other countries made similar claims. Charges were lodged by Tanzania, Nicaragua, Mexico, and the Philippines. The Gardasil vaccine claimed to protect against a genital 'virus' known as HPV has also been linked to infertility. Big Pharma and the WHO (same thing) are criminal and satanic entities. Then there's the Bill Gates Foundation which is connected through funding and shared interests with 20 pharmaceutical giants and laboratories. He stands accused of directing the policy of United Nations Children's Fund (UNICEF), vaccine alliance GAVI, and other groupings, to advance the vaccine agenda and silence opposition at great cost to women and children. At the same time Gates wants to reduce the global population. Coincidence?

Great Reset = Smart Grid = new human

The Cult agenda I have been exposing for 30 years is now being openly promoted by Cult assets like Gates and Klaus Schwab of the World Economic Forum under code-terms like the 'Great Reset', 'Build Back Better' and 'a rare but narrow window of opportunity to reflect, reimagine, and reset our world'. What provided this 'rare but narrow window of opportunity'? The 'Covid' hoax did. Who created that? *They* did. My books from not that long ago warned about the planned 'Internet of Things' (IoT) and its implications for human freedom. This was the plan to connect all technology to the Internet and artificial intelligence and today we are way down that road with an estimated 36 billion devices connected to the World Wide Web and that figure is projected to be 76 billion by 2025. I further warned that the Cult planned to go beyond that to the Internet of *Everything* when the human brain was connected via AI to the Internet and Kurzweil's 'cloud'. Now we have Cult operatives like Schwab calling for precisely that under the term 'Internet of Bodies', a fusion of the physical, digital and biological into one centrally-controlled Smart Grid system which the Cult refers to as the 'Fourth Industrial Revolution'. They talk about the 'biological', but they really mean the synthetic-biological which is required to fully integrate the human body and brain into the Smart Grid and artificial intelligence planned to replace the human mind. We have everything being synthetically manipulated including the natural world through GMO and smart dust, the food we eat and the human body itself with synthetic 'vaccines'. I said in *The Answer* that we would see the Cult push for synthetic meat to replace animals and in February, 2021, the so predictable psychopath Bill Gates called for the introduction of synthetic meat to save us all from 'climate change'. The climate hoax just keeps on giving like the 'Covid' hoax. The war on meat by vegan activists is a carbon (oops, sorry) copy of the manipulation of transgender activists. They have no idea (except their inner core) that they are being used to promote and impose the agenda of the Cult or that they are only the *vehicle* and not the *reason*. This is not to say those who choose not to eat meat shouldn't be respected and supported in that right, but there are ulterior motives

for those in power. A *Forbes* article in December, 2019, highlighted the plan so beloved of Schwab and the Cult under the heading: 'What Is The Internet of Bodies? And How Is It Changing Our World?' The article said the human body is the latest data platform (remember 'our vaccine is an operating system'). *Forbes* described the plan very accurately and the words could have come straight out of my books from long before:

The Internet of Bodies (IoB) is an extension of the IoT and basically connects the human body to a network through devices that are ingested, implanted, or connected to the body in some way. Once connected, data can be exchanged, and the body and device can be remotely monitored and controlled.

They were really describing a human hive mind with human perception centrally-dictated via an AI connection as well as allowing people to be 'remotely monitored and controlled'. Everything from a fridge to a human mind could be directed from a central point by these insane psychopaths and 'Covid vaccines' are crucial to this. Forbes explained the process I mentioned earlier of holdable and wearable technology followed by implantable. The article said there were three generations of the Internet of Bodies that include:

- Body external: These are wearable devices such as Apple Watches or Fitbits that can monitor our health.

- Body internal: These include pacemakers, cochlear implants, and digital pills that go inside our bodies to monitor or control various aspects of health.

- Body embedded: The third generation of the Internet of Bodies is embedded technology where technology and the human body are melded together and have a real-time connection to a remote machine.

Forbes noted the development of the Brain Computer Interface (BCI) which merges the brain with an external device for monitoring and controlling in real-time. 'The ultimate goal is to help restore function to individuals with disabilities by using brain signals rather than conventional neuromuscular pathways.' Oh, do fuck off. The goal of brain interface technology is controlling human thought and emotion from the central point in a hive mind serving its masters wishes. Many people are now agreeing to be chipped to open doors without a key. You can recognise them because they'll be wearing a mask, social distancing and lining up for the 'vaccine'. The Cult plans a Great Reset money system after they have completed the demolition of the global economy in which 'money' will be exchanged through communication with body operating systems. Rand Corporation, a Cult-owned think tank, said of the Internet of Bodies or IoB:

> Internet of Bodies technologies fall under the broader IoT umbrella. But as the name suggests, IoB devices introduce an even more intimate interplay between humans and gadgets. IoB devices monitor the human body, collect health metrics and other personal information, and transmit those data over the Internet. Many devices, such as fitness trackers, are already in use … IoB devices … and those in development can track, record, and store users' whereabouts, bodily functions, and what they see, hear, and even think.

Schwab's World Economic Forum, a long-winded way of saying 'fascism' or 'the Cult', has gone full-on with the Internet of Bodies in the 'Covid' era. 'We're entering the era of the Internet of Bodies', it declared, 'collecting our physical data via a range of devices that can be implanted, swallowed or worn'. The result would be a huge amount of health-related data that could improve human wellbeing around the world, and prove crucial in fighting the 'Covid-19 pandemic'. Does anyone think these clowns care about 'human wellbeing' after the death and devastation their pandemic hoax has purposely caused? Schwab and co say we should move forward with the Internet of Bodies because 'Keeping track of symptoms could help us stop the spread of infection, and quickly detect new cases'. How wonderful, but keeping track' is all they are really bothered

about. Researchers were investigating if data gathered from smartwatches and similar devices could be used as viral infection alerts by tracking the user's heart rate and breathing. Schwab said in his 2018 book *Shaping the Future of the Fourth Industrial Revolution*:

> The lines between technologies and beings are becoming blurred and not just by the ability to create lifelike robots or synthetics. Instead it is about the ability of new technologies to literally become part of us. Technologies already influence how we understand ourselves, how we think about each other, and how we determine our realities. As the technologies … give us deeper access to parts of ourselves, we may begin to integrate digital technologies into our bodies.

You can see what the game is. Twenty-four hour control and people – if you could still call them that – would never know when something would go ping and take them out of circulation. It's the most obvious rush to a global fascist dictatorship and the complete submission of humanity and yet still so many are locked away in their Cult-induced perceptual coma and can't see it.

Smart Grid control centres

The human body is being transformed by the 'vaccines' and in other ways into a synthetic cyborg that can be attached to the global Smart Grid which would be controlled from a central point and other sub-locations of Grid manipulation. Where are these planned to be? Well, China for a start which is one of the Cult's biggest centres of operation. The technological control system and technocratic rule was incubated here to be unleashed across the world after the 'Covid' hoax came out of China in 2020. Another Smart Grid location that will surprise people new to this is Israel. I have exposed in *The Trigger* how Sabbatian technocrats, intelligence and military operatives were behind the horrors of 9/11 and not '19 Arab hijackers' who somehow manifested the ability to pilot big passenger airliners when instructors at puddle-jumping flying schools described some of them as a joke. The 9/11 attacks were made possible through control of civilian and military air computer systems and those of the White House, Pentagon and connected agencies. See *The Trigger* – it

will blow your mind. The controlling and coordinating force were the Sabbatian networks in Israel and the United States which by then had infiltrated the entire US government, military and intelligence system. The real name of the American Deep State is 'Sabbatian State'. Israel is a tiny country of only nine million people, but it is one of the global centres of cyber operations and fast catching Silicon Valley in importance to the Cult. Israel is known as the 'start-up nation' for all the cyber companies spawned there with the Sabbatian specialisation of 'cyber security' that I mentioned earlier which gives those companies access to computer systems of their clients in real time through 'backdoors' written into the coding when security software is downloaded. The Sabbatian centre of cyber operations outside Silicon Valley is the Israeli military Cyber Intelligence Unit, the biggest infrastructure project in Israel's history, headquartered in the desert-city of Beersheba and involving some 20,000 'cyber soldiers'. Here are located a literal army of Internet trolls scanning social media, forums and comment lists for anyone challenging the Cult agenda. The UK military has something similar with its 77th Brigade and associated operations. The Beersheba complex includes research and development centres for other Cult operations such as Intel, Microsoft, IBM, Google, Apple, Hewlett-Packard, Cisco Systems, Facebook and Motorola. Techcrunch.com ran an article about the Beersheba global Internet technology centre headlined 'Israel's desert city of Beersheba is turning into a cybertech oasis':

> The military's massive relocation of its prestigious technology units, the presence of multinational and local companies, a close proximity to Ben Gurion University and generous government subsidies are turning Beersheba into a major global cybertech hub. Beersheba has all of the ingredients of a vibrant security technology ecosystem, including Ben Gurion University with its graduate program in cybersecurity and Cyber Security Research Center, and the presence of companies such as EMC, Deutsche Telekom, PayPal, Oracle, IBM, and Lockheed Martin. It's also the future home of the INCB (Israeli National Cyber Bureau); offers a special income tax incentive for cyber security companies, and was the site for the relocation of the army's intelligence corps units.

Sabbatians have taken over the cyber world through the following process: They scan the schools for likely cyber talent and develop them at Ben Gurion University and their period of conscription in the Israeli Defense Forces when they are stationed at the Beersheba complex. When the cyber talented officially leave the army they are funded to start cyber companies with technology developed by themselves or given to them by the state. Much of this is stolen through backdoors of computer systems around the world with America top of the list. Others are sent off to Silicon Valley to start companies or join the major ones and so we have many major positions filled by apparently 'Jewish' but really Sabbatian operatives. Google, YouTube and Facebook are all run by 'Jewish' CEOs while Twitter is all but run by ultra-Zionist hedge-fund shark Paul Singer. At the centre of the Sabbatian global cyber web is the Israeli army's Unit 8200 which specialises in hacking into computer systems of other countries, inserting viruses, gathering information, instigating malfunction, and even taking control of them from a distance. A long list of Sabbatians involved with 9/11, Silicon Valley and Israeli cyber security companies are operatives of Unit 8200. This is not about Israel. It's about the Cult. Israel is planned to be a Smart Grid hub as with China and what is happening at Beersheba is not for the benefit of Jewish people who are treated disgustingly by the Sabbatian elite that control the country. A glance at the Nuremberg Codes will tell you that.

The story is much bigger than 'Covid', important as that is to where we are being taken. Now, though, it's time to really strap in. There's more … much more …

CHAPTER ELEVEN

Who controls the Cult?

Awake, arise or be forever fall'n
John Milton, *Paradise Lost*

I have exposed this far the level of the Cult conspiracy that operates in the world of the seen and within the global secret society and satanic network which operates in the shadows one step back from the seen. The story, however, goes much deeper than that.

The 'Covid' hoax is major part of the Cult agenda, but only part, and to grasp the biggest picture we have to expand our attention beyond the realm of human sight and into the infinity of possibility that we cannot see. It is from here, ultimately, that humanity is being manipulated into a state of total control by the force which dictates the actions of the Cult. How much of reality can we see? Next to damn all is the answer. We may appear to see all there is to see in the 'space' our eyes survey and observe, but little could be further from the truth. The human 'world' is only a tiny band of frequency that the body's visual and perceptual systems can decode into *perception* of a 'world'. According to mainstream science the electromagnetic spectrum is 0.005 percent of what exists in the Universe (Fig 10). The maximum estimate I have seen is 0.5 percent and either way it's miniscule. I say it is far, far, smaller even than 0.005 percent when you compare reality we see with the totality of reality that we don't. Now get this if you are new to such information: Visible light, the only band of frequency that we can see, is a *fraction* of the 0.005

percent (Fig 11 overleaf). Take this further and realise that our universe is one of infinite universes and that universes are only a fragment of overall reality – *infinite* reality. Then compare that with the almost infinitesimal frequency band of visible light or human sight. You see that humans are as near blind as it is possible to be without actually being so. Artist and filmmaker, Sergio Toporek, said:

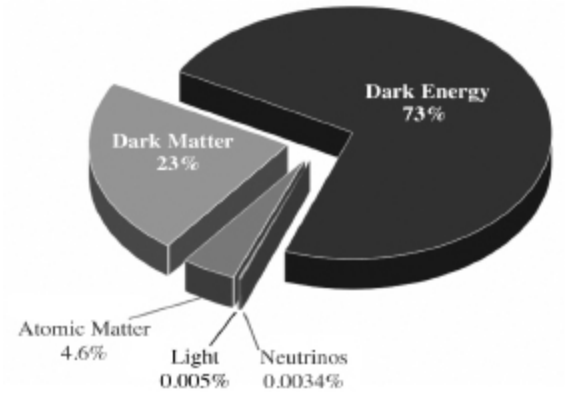


Figure 10: Humans can perceive such a tiny band of visual reality it's laughable.

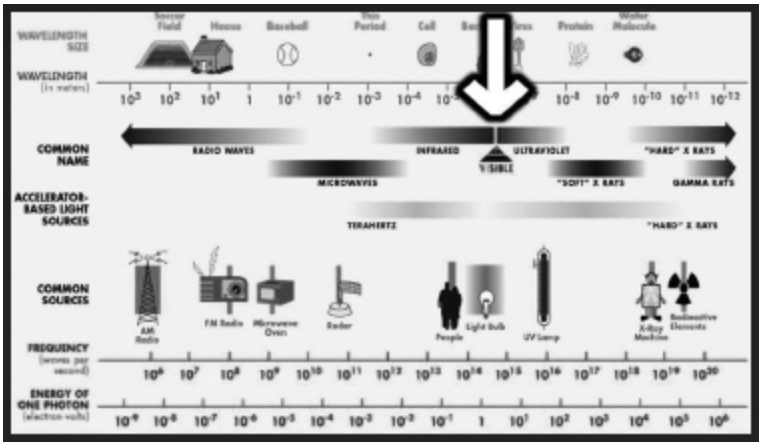


Figure 11: We can see a smear of the 0.005 percent electromagnetic spectrum, but we still know it all. Yep, makes sense.

Consider that you can see less than 1% of the electromagnetic spectrum and hear less than 1% of the acoustic spectrum. 90% of the cells in your body carry their own microbial DNA and are not 'you'. The atoms in your body are 99.9999999999999999% empty space and none of them are the ones you were born with ... Human beings have 46 chromosomes, two less than a potato.

The existence of the rainbow depends on the conical photoreceptors in your eyes; to animals without cones, the rainbow does not exist. So you don't just look at a rainbow, you create it. This is pretty amazing, especially considering that all the beautiful colours you see represent less than 1% of the electromagnetic spectrum.

Suddenly the 'world' of humans looks a very different place. Take into account, too, that Planet Earth when compared with the projected size of this single universe is the equivalent of a billionth of a pinhead. Imagine the ratio that would be when compared to infinite reality. To think that Christianity once insisted that Earth and humanity were the centre of everything. This background is vital if we are going to appreciate the nature of 'human' and how we can be manipulated by an unseen force. To human visual reality virtually *everything* is unseen and yet the prevailing perception within the institutions and so much of the public is that if we can't see it, touch it, hear it, taste it and smell it then it cannot exist. Such perception is indoctrinated and encouraged by the Cult and its agents because it isolates believers in the strictly limited, village-idiot, realm of the five senses where perceptions can be firewalled and information controlled. Most of those perpetuating the 'this-world-is-all-there-is' insanity are themselves indoctrinated into believing the same delusion. While major players and influencers know that official reality is laughable most of those in science, academia and medicine really believe the nonsense they peddle and teach succeeding generations. Those who challenge the orthodoxy are dismissed as nutters and freaks to protect the manufactured illusion from exposure. Observe the dynamic of the 'Covid' hoax and you will see how that takes the same form. The inner-circle psychopaths knows it's a gigantic scam, but almost the entirety of those imposing their fascist rules believe that 'Covid' is all that they're told it is.

Stolen identity

Ask people who they are and they will give you their name, place of birth, location, job, family background and life story. Yet that is not who they are – it is what they are *experiencing*. The difference is *absolutely crucial*. The true 'I', the eternal, infinite 'I', is consciousness,

a state of being aware. Forget 'form'. That is a vehicle for a brief experience. Consciousness does not come *from* the brain, but *through* the brain and even that is more symbolic than literal. We are awareness, pure awareness, and this is what withdraws from the body at what we call 'death' to continue our eternal beingness, *isness*, in other realms of reality within the limitlessness of infinity or the Biblical 'many mansions in my father's house'. Labels of a human life, man, woman, transgender, black, white, brown, nationality, circumstances and income are not who we are. They are what we are – awareness – is *experiencing* in a brief connection with a band of frequency we call 'human'. The labels are not the self; they are, to use the title of one of my books, a *Phantom* Self. I am not David Icke born in Leicester, England, on April 29th, 1952. I am the consciousness *having that experience*. The Cult and its non-human masters seek to convince us through the institutions of 'education', science, medicine, media and government that what we are *experiencing* is who we *are*. It's so easy to control and direct perception locked away in the bewildered illusions of the five senses with no expanded radar. Try, by contrast, doing the same with a humanity aware of its true self and its true power to consciously create its reality and experience. How is it possible to do this? We do it all day every day. If you perceive yourself as 'little me' with no power to impact upon your life and the world then your life experience will reflect that. You will hand the power you don't think you have to authority in all its forms which will use it to control your experience. This, in turn, will appear to confirm your perception of 'little me' in a self-fulfilling feedback loop. But that is what 'little me' really is – a *perception*. We are all 'big-me', infinite me, and the Cult has to make us forget that if its will is to prevail. We are therefore manipulated and pressured into self-identifying with human labels and not the consciousness/awareness *experiencing* those human labels.

The phenomenon of identity politics is a Cult-instigated manipulation technique to sub-divide previous labels into even smaller ones. A United States university employs this list of letters to

describe student identity: LGBTTQQFAGPBDSM or lesbian, gay, bisexual, transgender, transsexual, queer, questioning, flexual, asexual, gender-fuck, polyamorous, bondage/discipline, dominance/submission and sadism/masochism. I'm sure other lists are even longer by now as people feel the need to self-identity the 'I' with the minutiae of race and sexual preference. Wokers programmed by the Cult for generations believe this is about 'inclusivity' when it's really the Cult locking them away into smaller and smaller versions of Phantom Self while firewalling them from the influence of their true self, the infinite, eternal 'I'. You may notice that my philosophy which contends that we are all unique points of attention/awareness within the same infinite whole or Oneness is the ultimate non-racism. The very sense of Oneness makes the judgement of people by their body-type, colour or sexuality utterly ridiculous and confirms that racism has no understanding of reality (including anti-white racism). Yet despite my perception of life Cult agents and fast-asleep Wokers label me racist to discredit my information while they are themselves phenomenally racist and sexist. All they see is race and sexuality and they judge people as good or bad, demons or untouchables, by their race and sexuality. All they see is *Phantom Self* and perceive themselves in terms of Phantom Self. They are pawns and puppets of the Cult agenda to focus attention and self-identity in the five senses and play those identities against each other to divide and rule. Columbia University has introduced segregated graduations in another version of social distancing designed to drive people apart and teach them that different racial and cultural groups have nothing in common with each other. The last thing the Cult wants is unity. Again the pump-primers of this will be Cult operatives in the knowledge of what they are doing, but the rest are just the Phantom Self blind leading the Phantom Self blind. We *do* have something in common – we are all *the same consciousness* having different temporary experiences.

What is this 'human'?

Yes, what *is* 'human'? That is what we are supposed to be, right? I mean 'human'? True, but 'human' is the experience not the 'I'. Break it down to basics and 'human' is the way that information is processed. If we are to experience and interact with this band of frequency we call the 'world' we must have a vehicle that operates within that band of frequency. Our consciousness in its prime form cannot do that; it is way beyond the frequency of the human realm. My consciousness or awareness could not tap these keys and pick up the cup in front of me in the same way that radio station A cannot interact with radio station B when they are on different frequencies. The human body is the means through which we have that interaction. I have long described the body as a biological computer which processes information in a way that allows consciousness to experience this reality. The body is a receiver, transmitter and processor of information in a particular way that we call human. We visually perceive only the world of the five senses in a wakened state – that is the limit of the body's visual decoding system. In truth it's not even visual in the way we experience 'visual reality' as I will come to in a moment. We are 'human' because the body processes the information sources of human into a reality and behaviour system that we *perceive* as human. Why does an elephant act like an elephant and not like a human or a duck? The elephant's biological computer is a different information field and processes information according to that program into a visual and behaviour type we call an elephant. The same applies to everything in our reality. These body information fields are perpetuated through procreation (like making a copy of a software program). The Cult wants to break that cycle and intervene technologically to transform the human information field into one that will change what we call humanity. If it can change the human information field it will change the way that field processes information and change humanity both 'physically' and psychologically. Hence the *messenger* (information) RNA 'vaccines' and so much more that is targeting human genetics by changing the body's information – *messaging* – construct through food, drink, radiation, toxicity and other means.

Reality that we experience is nothing like reality as it really is in the same way that the reality people experience in virtual reality games is not the reality they are really living in. The game is only a decoded source of information that appears to be a reality. Our world is also an information construct – a *simulation* (more later). In its base form our reality is a wavefield of information much the same in theme as Wi-Fi. The five senses decode wavefield information into electrical information which they communicate to the brain to decode into holographic (illusory 'physical') information. Different parts of the brain specialise in decoding different senses and the information is fused into a reality that appears to be outside of us but is really inside the brain and the genetic structure in general (Fig 12 overleaf). DNA is a receiver-transmitter of information and a vital part of this decoding process and the body's connection to other realities. Change DNA and you change the way we decode and connect with reality – see 'Covid vaccines'. Think of computers decoding Wi-Fi. You have information encoded in a radiation field and the computer decodes that information into a very different form on the screen. You can't see the Wi-Fi until its information is made manifest on the screen and the information on the screen is inside the computer and not outside. I have just described how we decode the 'human world'. All five senses decode the waveform 'Wi-Fi' field into electrical signals and the brain (computer) constructs reality inside the brain and not outside – 'You don't just look at a rainbow, you create it'. Sound is a simple example. We don't hear sound until the brain decodes it. Waveform sound waves are picked up by the hearing sense and communicated to the brain in an electrical form to be decoded into the sounds that we hear. Everything we hear is inside the brain along with everything we see, feel, smell and taste. Words and language are waveform fields generated by our vocal chords which pass through this process until they are decoded by the brain into words that we hear. Different languages are different frequency fields or sound waves generated by vocal chords. Late British philosopher Alan Watts said:

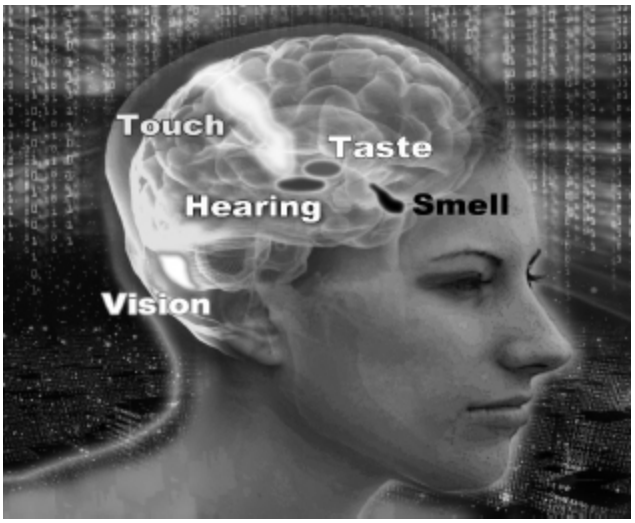


Figure 12: The brain receives information from the five senses and constructs from that our perceived reality.

[Without the brain] the world is devoid of light, heat, weight, solidity, motion, space, time or any other imaginable feature. All these phenomena are interactions, or transactions, of vibrations with a certain arrangement of neurons.

That's exactly what they are and scientist Robert Lanza describes in his book, *Biocentrism*, how we decode electromagnetic waves and energy into visual and 'physical' experience. He uses the example of a flame emitting photons, electromagnetic energy, each pulsing electrically and magnetically:

… these … invisible electromagnetic waves strike a human retina, and if (and only if) the waves happen to measure between 400 and 700 nano meters in length from crest to crest, then their energy is just right to deliver a stimulus to the 8 million cone-shaped cells in the retina.

Each in turn send an electrical pulse to a neighbour neuron, and on up the line this goes, at 250 mph, until it reaches the … occipital lobe of the brain, in the back of the head. There, a cascading complex of neurons fire from the incoming stimuli, and we subjectively perceive this experience as a yellow brightness occurring in a place we have been conditioned to call the 'external world'.

You hear what you decode

If a tree falls or a building collapses they make no noise unless someone is there to decode the energetic waves generated by the disturbance into what we call sound. Does a falling tree make a noise? Only if you hear it – *decode* it. Everything in our reality is a frequency field of information operating within the overall 'Wi-Fi' field that I call The Field. A vibrational disturbance is generated in The Field by the fields of the falling tree or building. These disturbance waves are what we decode into the sound of them falling. If no one is there to do that then neither will make any noise. Reality is created by the observer – *decoder* – and the *perceptions* of the observer affect the decoding process. For this reason different people – different *perceptions* – will perceive the same reality or situation in a different way. What one may perceive as a nightmare another will see as an opportunity. The question of why the Cult is so focused on controlling human perception now answers itself. All experienced reality is the act of decoding and we don't experience Wi-Fi until it is decoded on the computer screen. The sight and sound of an Internet video is encoded in the Wi-Fi all around us, but we don't see or hear it until the computer decodes that information. Taste, smell and touch are all phenomena of the brain as a result of the same process. We don't taste, smell or feel anything except in the brain and there are pain relief techniques that seek to block the signal from the site of discomfort to the brain because if the brain doesn't decode that signal we don't feel pain. Pain is in the brain and only appears to be at the point of impact thanks to the feedback loop between them. We don't see anything until electrical information from the sight senses is decoded in an area at the back of the brain. If that area is damaged we can go blind when our eyes are perfectly okay. So why do we go blind if we damage an eye? We damage the information processing between the waveform visual information and the visual decoding area of the brain. If information doesn't reach the brain in a form it can decode then we can't see the visual reality that it represents. What's more the brain is decoding only a fraction of the information it receives and the rest is absorbed by the

sub-conscious mind. This explanation is from the science magazine, *Wonderpedia*:

> Every second, 11 million sensations crackle along these [brain] pathways ... The brain is confronted with an alarming array of images, sounds and smells which it rigorously filters down until it is left with a manageable list of around 40. Thus 40 sensations per second make up what we perceive as reality.

The 'world' is not what people are told to believe that is it and the inner circles of the Cult *know that*.

Illusory 'physical' reality

We can only see a smear of 0.005 percent of the Universe which is only one of a vast array of universes – 'mansions' – within infinite reality. Even then the brain decodes only 40 pieces of information ('sensations') from a potential *11 million* that we receive every second. Two points strike you from this immediately: The sheer breathtaking stupidity of believing we know anything so rigidly that there's nothing more to know; and the potential for these processes to be manipulated by a malevolent force to control the reality of the population. One thing I can say for sure with no risk of contradiction is that when you can perceive an almost indescribable fraction of infinite reality there is always more to know as in tidal waves of it. Ancient Greek philosopher Socrates was so right when he said that wisdom is to know how little we know. How obviously true that is when you think that we are experiencing a physical world of solidity that is neither physical nor solid and a world of apartness when everything is connected. Cult-controlled 'science' dismisses the so-called 'paranormal' and all phenomena related to that when the 'para'-normal is perfectly normal and explains the alleged 'great mysteries' which dumbfound scientific minds. There is a reason for this. A 'scientific mind' in terms of the mainstream is a material mind, a five-sense mind imprisoned in see it, touch it, hear it, smell it and taste it. Phenomena and happenings that can't be explained that way leave the 'scientific mind' bewildered and the rule is that if they

can't account for why something is happening then it can't, by definition, be happening. I beg to differ. Telepathy is thought waves passing through The Field (think wave disturbance again) to be decoded by someone able to connect with that wavelength (information). For example: You can pick up the thought waves of a friend at any distance and at the very least that will bring them to mind. A few minutes later the friend calls you. 'My god', you say, 'that's incredible – I was just thinking of you.' Ah, but *they* were thinking of *you* before they made the call and that's what you decoded. Native peoples not entrapped in five-sense reality do this so well it became known as the 'bush telegraph'. Those known as psychics and mediums (genuine ones) are doing the same only across dimensions of reality. 'Mind over matter' comes from the fact that matter and mind are the *same*. The state of one influences the state of the other. Indeed one *and* the other are illusions. They are aspects of the same field. Paranormal phenomena are all explainable so why are they still considered 'mysteries' or not happening? Once you go down this road of understanding you begin to expand awareness beyond the five senses and that's the nightmare for the Cult.

Figure 13: Holograms are not solid, but the best ones appear to be.

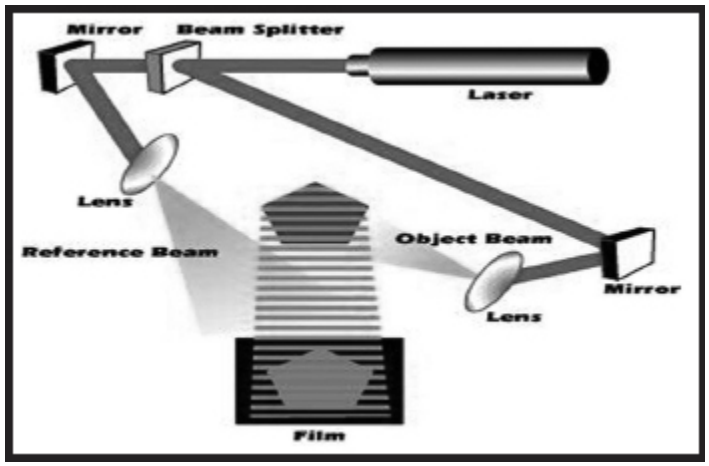


Figure 14: How holograms are created by capturing a waveform version of the subject image.

Holographic 'solidity'

Our reality is not solid, it is holographic. We are now well aware of holograms which are widely used today. Two-dimensional information is decoded into a three-dimensional reality that is not solid although can very much appear to be (Fig 13). Holograms are created with a laser divided into two parts. One goes directly onto a holographic photographic print ('reference beam') and the other takes a waveform image of the subject ('working beam') before being directed onto the print where it 'collides' with the other half of the laser (Fig 14). This creates a *waveform* interference pattern which contains the wavefield information of whatever is being photographed (Fig 15 overleaf). The process can be likened to dropping pebbles in a pond. Waves generated by each one spread out across the water to collide with the others and create a wave representation of where the stones fell and at what speed, weight and distance. A waveform interference pattern of a hologram is akin to the waveform information in The Field which the five senses decode into electrical signals to be decoded by the brain into a holographic illusory 'physical' reality. In the same way when a laser (think human attention) is directed at the waveform interference pattern a three-dimensional version of the subject is projected into apparently 'solid' reality (Fig 16). An amazing trait of holograms reveals more 'paranormal mysteries'. Information of the *whole*

hologram is encoded in waveform in every part of the interference pattern by the way they are created. This means that every *part* of a hologram is a smaller version of the whole. Cut the interference wave-pattern into four and you won't get four parts of the image. You get quarter-sized versions of the *whole* image. The body is a hologram and the same applies. Here we have the basis of acupuncture, reflexology and other forms of healing which identify representations of the whole body in all of the parts, hands, feet, ears, everywhere. Skilled palm readers can do what they do because the information of whole body is encoded in the hand. The concept of as above, so below, comes from this.

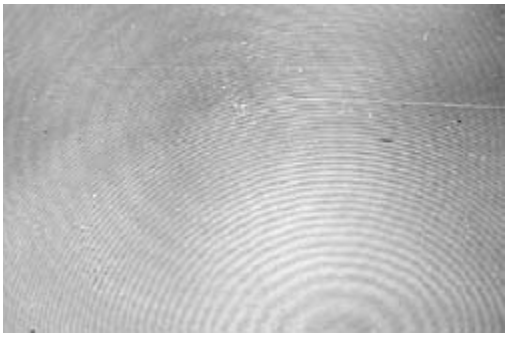


Figure 15: A waveform interference pattern that holds the information that transforms into a hologram.

Figure 16: Holographic people including 'Elvis' holographically inserted to sing a duet with Celine Dion.

The question will be asked of why, if solidity is illusory, we can't just walk through walls and each other. The resistance is not solid against solid; it is electromagnetic field against electromagnetic field and we decode this into the *experience* of solid against solid. We should also not underestimate the power of belief to dictate reality. What you believe is impossible *will be*. Your belief impacts on your decoding processes and they won't decode what you think is impossible. What we believe we perceive and what we perceive we experience. 'Can't dos' and 'impossibles' are like a firewall in a computer system that won't put on the screen what the firewall blocks. How vital that is to understanding how human experience has been hijacked. I explain in *The Answer*, *Everything You Need To Know But Have Never Been Told* and other books a long list of 'mysteries' and 'paranormal' phenomena that are not mysterious and perfectly normal once you realise what reality is and how it works. 'Ghosts' can be seen to pass through 'solid' walls because the walls are not solid and the ghost is a discarnate entity operating on a frequency so different to that of the wall that it's like two radio stations sharing the same space while never interfering with each other. I have seen ghosts do this myself. The apartness of people and objects is also an illusion. Everything is connected by the Field like all sea life is connected by the sea. It's just that within the limits of our visual reality we only 'see' holographic information and not the field of information that connects everything and from which the holographic world is made manifest. If you can only see holographic 'objects' and not the field that connects them they will appear to you as unconnected to each other in the same way that we see the computer while not seeing the Wi-Fi.

What you don't know *can* hurt you

Okay, we return to those 'two worlds' of human society and the Cult with its global network of interconnecting secret societies and satanic groups which manipulate through governments, corporations, media, religions, etc. The fundamental difference between them is *knowledge*. The idea has been to keep humanity

ignorant of the plan for its total enslavement underpinned by a crucial ignorance of reality – who we are and where we are – and how we interact with it. 'Human' should be the interaction between our expanded eternal consciousness and the five-sense body experience. We are meant to be *in* this world in terms of the five senses but not *of* this world in relation to our greater consciousness and perspective. In that state we experience the small picture of the five senses within the wider context of the big picture of awareness beyond the five senses. Put another way the five senses see the dots and expanded awareness connects them into pictures and patterns that give context to the apparently random and unconnected. Without the context of expanded awareness the five senses see only apartness and randomness with apparently no meaning. The Cult and its other-dimensional controllers seek to intervene in the frequency realm where five-sense reality is supposed to connect with expanded reality and to keep the two apart (more on this in the final chapter). When that happens five-sense mental and emotional processes are no longer influenced by expanded awareness, or the True 'I', and instead are driven by the isolated perceptions of the body's decoding systems. They are in the world *and* of it. Here we have the human plight and why humanity with its potential for infinite awareness can be so easily manipulatable and descend into such extremes of stupidity.

Once the Cult isolates five-sense mind from expanded awareness it can then program the mind with perceptions and beliefs by controlling information that the mind receives through the 'education' system of the formative years and the media perceptual bombardment and censorship of an entire lifetime. Limit perception and a sense of the possible through limiting knowledge by limiting and skewing information while censoring and discrediting that which could set people free. As the title of another of my books says *… And The Truth Shall Set You Free*. For this reason the last thing the Cult wants in circulation is the truth about anything – especially the reality of the eternal 'I' – and that's why it is desperate to control information. The Cult knows that information becomes perception

which becomes behaviour which, collectively, becomes human society. Cult-controlled and funded mainstream 'science' denies the existence of an eternal 'I' and seeks to dismiss and trash all evidence to the contrary. Cult-controlled mainstream religion has a version of 'God' that is little more than a system of control and dictatorship that employs threats of damnation in an afterlife to control perceptions and behaviour in the here and now through fear and guilt. Neither is true and it's the 'neither' that the Cult wishes to suppress. This 'neither' is that everything is an expression, a point of attention, within an infinite state of consciousness which is the real meaning of the term 'God'.

Perceptual obsession with the 'physical body' and five-senses means that 'God' becomes personified as a bearded bloke sitting among the clouds or a raging bully who loves us if we do what 'he' wants and condemns us to the fires of hell if we don't. These are no more than a 'spiritual' fairy tales to control and dictate events and behaviour through fear of this 'God' which has bizarrely made 'God-fearing' in religious circles a state to be desired. I would suggest that fearing *anything* is not to be encouraged and celebrated, but rather deleted. You can see why 'God fearing' is so beneficial to the Cult and its religions when *they* decide what 'God' wants and what 'God' demands (the Cult demands) that everyone do. As the great American comedian Bill Hicks said satirising a Christian zealot: 'I think what God meant to say.' How much of this infinite awareness ('God') that we access is decided by how far we choose to expand our perceptions, self-identity and sense of the possible. The scale of self-identity reflects itself in the scale of awareness that we can connect with and are influenced by – how much knowing and insight we have instead of programmed perception. You cannot expand your awareness into the infinity of possibility when you believe that you are little me Peter the postman or Mary in marketing and nothing more. I'll deal with this in the concluding chapter because it's crucial to how we turnaround current events.

Where the Cult came from

When I realised in the early 1990s there was a Cult network behind global events I asked the obvious question: When did it start? I took it back to ancient Rome and Egypt and on to Babylon and Sumer in Mesopotamia, the 'Land Between Two Rivers', in what we now call Iraq. The two rivers are the Tigris and Euphrates and this region is of immense historical and other importance to the Cult, as is the land called Israel only 550 miles away by air. There is much more going with deep esoteric meaning across this whole region. It's not only about 'wars for oil'. Priceless artefacts from Mesopotamia were stolen or destroyed after the American and British invasion of Iraq in 2003 justified by the lies of Boy Bush and Tony Blair (their Cult masters) about non-existent 'weapons of mass destruction'. Mesopotamia was the location of Sumer (about 5,400BC to 1,750BC), and Babylon (about 2,350BC to 539BC). Sabbatians may have become immensely influential in the Cult in modern times but they are part of a network that goes back into the mists of history. Sumer is said by historians to be the 'cradle of civilisation'. I disagree. I say it was the re-start of what we call human civilisation after cataclysmic events symbolised in part as the 'Great Flood' destroyed the world that existed before. These fantastic upheavals that I have been describing in detail in the books since the early1990s appear in accounts and legends of ancient cultures across the world and they are supported by geological and biological evidence. Stone tablets found in Iraq detailing the Sumer period say the cataclysms were caused by non-human 'gods' they call the Anunnaki. These are described in terms of extraterrestrial visitations in which knowledge supplied by the Anunnaki is said to have been the source of at least one of the world's oldest writing systems and developments in astronomy, mathematics and architecture that were way ahead of their time. I have covered this subject at length in *The Biggest Secret* and *Children of the Matrix* and the same basic 'Anunnaki' story can be found in Zulu accounts in South Africa where the late and very great Zulu high shaman Credo Mutwa told me that the Sumerian Anunnaki were known by Zulus as the Chitauri or 'children of the serpent'. See my six-hour video interview with Credo on this subject entitled *The*

Reptilian Agenda recorded at his then home near Johannesburg in 1999 which you can watch on the Ickonic media platform.

The Cult emerged out of Sumer, Babylon and Egypt (and elsewhere) and established the Roman Empire before expanding with the Romans into northern Europe from where many empires were savagely imposed in the form of Cult-controlled societies all over the world. Mass death and destruction was their calling card. The Cult established its centre of operations in Europe and European Empires were Cult empires which allowed it to expand into a global force. Spanish and Portuguese colonialists headed for Central and South America while the British and French targeted North America. Africa was colonised by Britain, France, Belgium, the Netherlands, Portugal, Spain, Italy, and Germany. Some like Britain and France moved in on the Middle East. The British Empire was by far the biggest for a simple reason. By now Britain was the headquarters of the Cult from which it expanded to form Canada, the United States, Australia and New Zealand. The Sun never set on the British Empire such was the scale of its occupation. London remains a global centre for the Cult along with Rome and the Vatican although others have emerged in Israel and China. It is no accident that the 'virus' is alleged to have come out of China while Italy was chosen as the means to terrify the Western population into compliance with 'Covid' fascism. Nor that Israel has led the world in 'Covid' fascism and mass 'vaccination'.

You would think that I would mention the United States here, but while it has been an important means of imposing the Cult's will it is less significant than would appear and is currently in the process of having what power it does have deleted. The Cult in Europe has mostly loaded the guns for the US to fire. America has been controlled from Europe from the start through Cult operatives in Britain and Europe. The American Revolution was an illusion to make it appear that America was governing itself while very different forces were pulling the strings in the form of Cult families such as the Rothschilds through the Rockefellers and other subordinates. The Rockefellers are extremely close to Bill Gates and

established both scalpel and drug 'medicine' and the World Health Organization. They play a major role in the development and circulation of vaccines through the Rockefeller Foundation on which Bill Gates said his Foundation is based. Why wouldn't this be the case when the Rockefellers and Gates are on the same team? Cult infiltration of human society goes way back into what we call history and has been constantly expanding and centralising power with the goal of establishing a global structure to dictate everything. Look how this has been advanced in great leaps with the 'Covid' hoax.

The non-human dimension

I researched and observed the comings and goings of Cult operatives through the centuries and even thousands of years as they were born, worked to promote the agenda within the secret society and satanic networks, and then died for others to replace them. Clearly there had to be a coordinating force that spanned this entire period while operatives who would not have seen the end goal in their lifetimes came and went advancing the plan over millennia. I went in search of that coordinating force with the usual support from the extraordinary synchronicity of my life which has been an almost daily experience since 1990. I saw common themes in religious texts and ancient cultures about a non-human force manipulating human society from the hidden. Christianity calls this force Satan, the Devil and demons; Islam refers to the Jinn or Djinn; Zulus have their Chitauri (spelt in other ways in different parts of Africa); and the Gnostic people in Egypt in the period around and before 400AD referred to this phenomena as the 'Archons', a word meaning rulers in Greek. Central American cultures speak of the 'Predators' among other names and the same theme is everywhere. I will use 'Archons' as a collective name for all of them. When you see how their nature and behaviour is described all these different sources are clearly talking about the same force. Gnostics described the Archons in terms of 'luminous fire' while Islam relates the Jinn to 'smokeless fire'. Some refer to beings in form that could occasionally be seen, but the most common of common theme is that they operate from

unseen realms which means almost all existence to the visual processes of humans. I had concluded that this was indeed the foundation of human control and that the Cult was operating within the human frequency band on behalf of this hidden force when I came across the writings of Gnostics which supported my conclusions in the most extraordinary way.

A sealed earthen jar was found in 1945 near the town of Nag Hammadi about 75-80 miles north of Luxor on the banks of the River Nile in Egypt. Inside was a treasure trove of manuscripts and texts left by the Gnostic people some 1,600 years earlier. They included 13 leather-bound papyrus codices (manuscripts) and more than 50 texts written in Coptic Egyptian estimated to have been hidden in the jar in the period of 400AD although the source of the information goes back much further. Gnostics oversaw the Great or Royal Library of Alexandria, the fantastic depository of ancient texts detailing advanced knowledge and accounts of human history. The Library was dismantled and destroyed in stages over a long period with the death-blow delivered by the Cult-established Roman Church in the period around 415AD. The Church of Rome was the Church of Babylon relocated as I said earlier. Gnostics were not a race. They were a way of perceiving reality. Whenever they established themselves and their information circulated the terrorists of the Church of Rome would target them for destruction. This happened with the Great Library and with the Gnostic Cathars who were burned to death by the psychopaths after a long period of oppression at the siege of the Castle of Monségur in southern France in 1244. The Church has always been terrified of Gnostic information which demolishes the official Christian narrative although there is much in the Bible that supports the Gnostic view if you read it in another way. To anyone studying the texts of what became known as the Nag Hammadi Library it is clear that great swathes of Christian and Biblical belief has its origin with Gnostics sources going back to Sumer. Gnostic themes have been twisted to manipulate the perceived reality of Bible believers. Biblical texts have been in the open for centuries where they could be changed while Gnostic

documents found at Nag Hammadi were sealed away and untouched for 1,600 years. What you see is what they wrote.

Use your *pneuma* not your *nous*

Gnosticism and Gnostic come from 'gnosis' which means knowledge, or rather *secret* knowledge, in the sense of spiritual awareness – knowledge about reality and life itself. The desperation of the Cult's Church of Rome to destroy the Gnostics can be understood when the knowledge they were circulating was the last thing the Cult wanted the population to know. Sixteen hundred years later the same Cult is working hard to undermine and silence me for the same reason. The dynamic between knowledge and ignorance is a constant. 'Time' appears to move on, but essential themes remain the same. We are told to 'use your nous', a Gnostic word for head/brain/intelligence. They said, however, that spiritual awakening or 'salvation' could only be secured by expanding awareness *beyond* what they called *nous* and into *pneuma* or Infinite Self. Obviously as I read these texts the parallels with what I have been saying since 1990 were fascinating to me. There is a universal truth that spans human history and in that case why wouldn't we be talking the same language 16 centuries apart? When you free yourself from the perception program of the five senses and explore expanded realms of consciousness you are going to connect with the same information no matter what the perceived 'era' within a manufactured timeline of a single and tiny range of manipulated frequency. Humans working with 'smart' technology or knocking rocks together in caves is only a timeline appearing to operate within the human frequency band. Expanded awareness and the knowledge it holds have always been there whether the era be Stone Age or computer age. We can only access that knowledge by opening ourselves to its frequency which the five-sense prison cell is designed to stop us doing. Gates, Fauci, Whitty, Vallance, Zuckerberg, Brin, Page, Wojcicki, Bezos, and all the others behind the 'Covid' hoax clearly have a long wait before their range of frequency can make that connection given that an open heart is

crucial to that as we shall see. Instead of accessing knowledge directly through expanded awareness it is given to Cult operatives by the secret society networks of the Cult where it has been passed on over thousands of years outside the public arena. Expanded realms of consciousness is where great artists, composers and writers find their inspiration and where truth awaits anyone open enough to connect with it. We need to go there fast.

Archon hijack

A fifth of the Nag Hammadi texts describe the existence and manipulation of the Archons led by a 'Chief Archon' they call 'Yaldabaoth', or the 'Demiurge', and this is the Christian 'Devil', 'Satan', 'Lucifer', and his demons. Archons in Biblical symbolism are the 'fallen ones' which are also referred to as fallen angels after the angels expelled from heaven according to the Abrahamic religions of Judaism, Christianity and Islam. These angels are claimed to tempt humans to 'sin' ongoing and you will see how accurate that symbolism is during the rest of the book. The theme of 'original sin' is related to the 'Fall' when Adam and Eve were 'tempted by the serpent' and fell from a state of innocence and 'obedience' (connection) with God into a state of disobedience (disconnection). The Fall is said to have brought sin into the world and corrupted everything including human nature. Yaldabaoth, the 'Lord Archon', is described by Gnostics as a 'counterfeit spirit', 'The Blind One', 'The Blind God', and 'The Foolish One'. The Jewish name for Yaldabaoth in Talmudic writings is Samael which translates as 'Poison of God', or 'Blindness of God'. You see the parallels. Yaldabaoth in Islamic belief is the Muslim Jinn devil known as Shaytan – Shaytan is Satan as the same themes are found all over the world in every religion and culture. The 'Lord God' of the Old Testament is the 'Lord Archon' of Gnostic manuscripts and that's why he's such a bloodthirsty bastard. Satan is known by Christians as 'the Demon of Demons' and Gnostics called Yaldabaoth the 'Archon of Archons'. Both are known as 'The Deceiver'. We are talking about the same 'bloke' for sure and these common themes

using different names, storylines and symbolism tell a common tale of the human plight.

Archons are referred to in Nag Hammadi documents as mind parasites, inverters, guards, gatekeepers, detainers, judges, pitiless ones and deceivers. The 'Covid' hoax alone is a glaring example of all these things. The Biblical 'God' is so different in the Old and New Testaments because they are not describing the same phenomenon. The vindictive, angry, hate-filled, 'God' of the Old Testament, known as Yahweh, is Yaldabaoth who is depicted in Cult-dictated popular culture as the 'Dark Lord', 'Lord of Time', Lord (Darth) Vader and Dormammu, the evil ruler of the 'Dark Dimension' trying to take over the 'Earth Dimension' in the Marvel comic movie, *Dr Strange*. Yaldabaoth is both the Old Testament 'god' and the Biblical 'Satan'. Gnostics referred to Yaldabaoth as the 'Great Architect of the Universe'and the Cult-controlled Freemason network calls their god 'the 'Great Architect of the Universe' (also Grand Architect). The 'Great Architect' Yaldabaoth is symbolised by the Cult as the all-seeing eye at the top of the pyramid on the Great Seal of the United States and the dollar bill. Archon is encoded in *arch*-itect as it is in *arch*-angels and *arch*-bishops. All religions have the theme of a force for good and force for evil in some sort of spiritual war and there is a reason for that – the theme is true. The Cult and its non-human masters are quite happy for this to circulate. They present themselves as the force for good fighting evil when they are really the force of evil (absence of love). The whole foundation of Cult modus operandi is inversion. They promote themselves as a force for good and anyone challenging them in pursuit of peace, love, fairness, truth and justice is condemned as a satanic force for evil. This has been the game plan throughout history whether the Church of Rome inquisitions of non-believers or 'conspiracy theorists' and 'anti-vaxxers' of today. The technique is the same whatever the timeline era.

Yaldabaoth is revolting (true)

Yaldabaoth and the Archons are said to have revolted against God with Yaldabaoth claiming to *be* God – the *All That Is*. The Old Testament 'God' (Yaldabaoth) demanded to be worshipped as such: ' *I am* the LORD, and there is none else, there is no God beside me' (Isaiah 45:5). I have quoted in other books a man who said he was the unofficial son of the late Baron Philippe de Rothschild of the Mouton-Rothschild wine producing estates in France who died in 1988 and he told me about the Rothschild 'revolt from God'. The man said he was given the name Phillip Eugene de Rothschild and we shared long correspondence many years ago while he was living under another identity. He said that he was conceived through 'occult incest' which (within the Cult) was 'normal and to be admired'. 'Phillip' told me about his experience attending satanic rituals with rich and famous people whom he names and you can see them and the wider background to Cult Satanism in my other books starting with *The Biggest Secret*. Cult rituals are interactions with Archontic 'gods'. 'Phillip' described Baron Philippe de Rothschild as 'a master Satanist and hater of God' and he used the same term 'revolt from God' associated with Yaldabaoth/Satan/Lucifer/the Devil in describing the Sabbatian Rothschild dynasty. 'I played a key role in my family's revolt from God', he said. That role was to infiltrate in classic Sabbatian style the Christian Church, but eventually he escaped the mind-prison to live another life. The Cult has been targeting religion in a plan to make worship of the Archons the global one-world religion. Infiltration of Satanism into modern 'culture', especially among the young, through music videos, stage shows and other means, is all part of this.

Nag Hammadi texts describe Yaldabaoth and the Archons in their prime form as energy – consciousness – and say they can take form if they choose in the same way that consciousness takes form as a human. Yaldabaoth is called 'formless' and represents a deeply inverted, distorted and chaotic state of consciousness which seeks to attached to humans and turn them into a likeness of itself in an attempt at assimilation. For that to happen it has to manipulate

humans into low frequency mental and emotional states that match its own. Archons can certainly appear in human form and this is the origin of the psychopathic personality. The energetic distortion Gnostics called Yaldabaoth *is* psychopathy. When psychopathic Archons take human form that human will be a psychopath as an expression of Yaldabaoth consciousness. Cult psychopaths are Archons in human form. The principle is the same as that portrayed in the 2009 *Avatar* movie when the American military travelled to a fictional Earth-like moon called Pandora in the Alpha Centauri star system to infiltrate a society of blue people, or Na'vi, by hiding within bodies that looked like the Na'vi. Archons posing as humans have a particular hybrid information field, part human, part Archon, (the ancient 'demigods') which processes information in a way that manifests behaviour to match their psychopathic evil, lack of empathy and compassion, and stops them being influenced by the empathy, compassion and love that a fully-human information field is capable of expressing. Cult bloodlines interbreed, be they royalty or dark suits, for this reason and you have their obsession with incest. Interbreeding with full-blown humans would dilute the Archontic energy field that guarantees psychopathy in its representatives in the human realm.

Gnostic writings say the main non-human forms that Archons take are *serpentine* (what I have called for decades 'reptilian' amid unbounded ridicule from the Archontically-programmed) and what Gnostics describe as 'an unborn baby or foetus with grey skin and dark, unmoving eyes'. This is an excellent representation of the ET 'Greys' of UFO folklore which large numbers of people claim to have seen and been abducted by – Zulu shaman Credo Mutwa among them. I agree with those that believe in extraterrestrial or interdimensional visitations today and for thousands of years past. No wonder with their advanced knowledge and technological capability they were perceived and worshipped as gods for technological and other 'miracles' they appeared to perform. Imagine someone arriving in a culture disconnected from the modern world with a smartphone and computer. They would be

seen as a 'god' capable of 'miracles'. The Renegade Mind, however, wants to know the source of everything and not only the way that source manifests as human or non-human. In the same way that a Renegade Mind seeks the original source material for the 'Covid virus' to see if what is claimed is true. The original source of Archons in form is consciousness – the distorted state of consciousness known to Gnostics as Yaldabaoth.

'Revolt from God' is energetic disconnection

Where I am going next will make a lot of sense of religious texts and ancient legends relating to 'Satan', Lucifer' and the 'gods'. Gnostic descriptions sync perfectly with the themes of my own research over the years in how they describe a consciousness distortion seeking to impose itself on human consciousness. I've referred to the core of infinite awareness in previous books as Infinite Awareness in Awareness of Itself. By that I mean a level of awareness that knows that it is all awareness and is aware of all awareness. From here comes the frequency of love in its true sense and balance which is what love is on one level – the balance of all forces into a single whole called Oneness and Isness. The more we disconnect from this state of love that many call 'God' the constituent parts of that Oneness start to unravel and express themselves as a part and not a whole. They become individualised as intellect, mind, selfishness, hatred, envy, desire for power over others, and such like. This is not a problem in the greater scheme in that 'God', the *All That Is*, can experience all these possibilities through different expressions of itself including humans. What we as expressions of the whole experience the *All That Is* experiences. We are the *All That Is* experiencing itself. As we withdraw from that state of Oneness we disconnect from its influence and things can get very unpleasant and very stupid. Archontic consciousness is at the extreme end of that. It has so disconnected from the influence of Oneness that it has become an inversion of unity and love, an inversion of everything, an inversion of life itself. Evil is appropriately live written backwards. Archontic consciousness is obsessed with death, an inversion of life,

and so its manifestations in Satanism are obsessed with death. They use inverted symbols in their rituals such as the inverted pentagram and cross. Sabbatians as Archontic consciousness incarnate invert Judaism and every other religion and culture they infiltrate. They seek disunity and chaos and they fear unity and harmony as they fear love like garlic to a vampire. As a result the Cult, Archons incarnate, act with such evil, psychopathy and lack of empathy and compassion disconnected as they are from the source of love. How could Bill Gates and the rest of the Archontic psychopaths do what they have to human society in the 'Covid' era with all the death, suffering and destruction involved and have no emotional consequence for the impact on others? Now you know. Why have Zuckerberg, Brin, Page, Wojcicki and company callously censored information warning about the dangers of the 'vaccine' while thousands have been dying and having severe, sometimes life-changing reactions? Now you know. Why have Tedros, Fauci, Whitty, Vallance and their like around the world been using case and death figures they're aware are fraudulent to justify lockdowns and all the deaths and destroyed lives that have come from that? Now you know. Why did Christian Drosten produce and promote a 'testing' protocol that he knew couldn't test for infectious disease which led to a global human catastrophe. Now you know. The Archontic mind doesn't give a shit (Fig 17). I personally think that Gates and major Cult insiders are a form of AI cyborg that the Archons want humans to become.

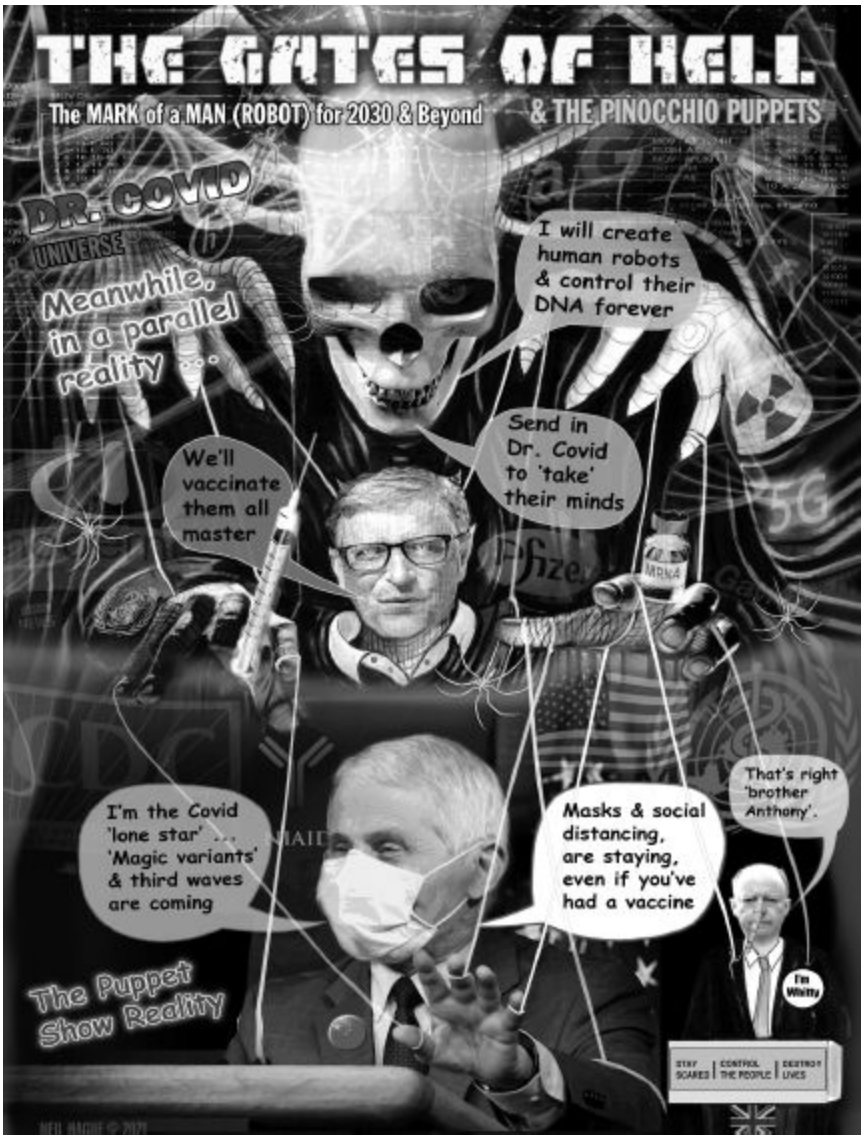


Figure 17: Artist Neil Hague's version of the 'Covid' hierarchy.

Human batteries

A state of such inversion does have its consequences, however. The level of disconnection from the Source of All means that you withdraw from that source of energetic sustenance and creativity. This means that you have to find your own supply of energetic power and it has – *us*. When the Morpheus character in the first *Matrix* movie held up a battery he spoke a profound truth when he said: 'The Matrix is a computer-generated dream world built to keep us under control in order to change the human being into one of

these.' The statement was true in all respects. We do live in a technologically-generated virtual reality simulation (more very shortly) and we have been manipulated to be an energy source for Archontic consciousness. The Disney-Pixar animated movie *Monsters, Inc*. in 2001 symbolised the dynamic when monsters in their world had no energy source and they would enter the human world to terrify children in their beds, catch the child's scream, terror (low-vibrational frequencies), and take that energy back to power the monster world. The lead character you might remember was a single giant eye and the symbolism of the Cult's all-seeing eye was obvious. Every thought and emotion is broadcast as a frequency unique to that thought and emotion. Feelings of love and joy, empathy and compassion, are high, quick, frequencies while fear, depression, anxiety, suffering and hate are low, slow, dense frequencies. Which kind do you think Archontic consciousness can connect with and absorb? In such a low and dense frequency state there's no way it can connect with the energy of love and joy. Archons can only feed off energy compatible with their own frequency and they and their Cult agents want to delete the human world of love and joy and manipulate the transmission of low vibrational frequencies through low-vibrational human mental and emotional states. *We are their energy source*. Wars are energetic banquets to the Archons – a world war even more so – and think how much low-frequency mental and emotional energy has been generated from the consequences for humanity of the 'Covid' hoax orchestrated by Archons incarnate like Gates.

The ancient practice of human sacrifice 'to the gods', continued in secret today by the Cult, is based on the same principle. 'The gods' are Archontic consciousness in different forms and the sacrifice is induced into a state of intense terror to generate the energy the Archontic frequency can absorb. Incarnate Archons in the ritual drink the blood which contains an adrenaline they crave which floods into the bloodstream when people are terrorised. Most of the sacrifices, ancient and modern, are children and the theme of 'sacrificing young virgins to the gods' is just code for children. They

have a particular pre-puberty energy that Archons want more than anything and the energy of the young in general is their target. The California Department of Education wants students to chant the names of Aztec gods (Archontic gods) once worshipped in human sacrifice rituals in a curriculum designed to encourage them to 'challenge racist, bigoted, discriminatory, imperialist/colonial beliefs', join 'social movements that struggle for social justice', and 'build new possibilities for a post-racist, post-systemic racism society'. It's the usual Woke crap that inverts racism and calls it anti-racism. In this case solidarity with 'indigenous tribes' is being used as an excuse to chant the names of 'gods' to which people were sacrificed (and still are in secret). What an example of Woke's inability to see beyond black and white, us and them, They condemn the colonisation of these tribal cultures by Europeans (quite right), but those cultures sacrificing people including children to their 'gods', and mass murdering untold numbers as the Aztecs did, is just fine. One chant is to the Aztec god Tezcatlipoca who had a man sacrificed to him in the 5th month of the Aztec calendar. His heart was cut out and he was eaten. Oh, that's okay then. Come on children … after three … Other sacrificial 'gods' for the young to chant their allegiance include Quetzalcoatl, Huitzilopochtli and Xipe Totec. The curriculum says that 'chants, affirmations, and energizers can be used to bring the class together, build unity around ethnic studies principles and values, and to reinvigorate the class following a lesson that may be emotionally taxing or even when student engagement may appear to be low'. Well, that's the cover story, anyway. Chanting and mantras are the repetition of a particular frequency generated from the vocal cords and chanting the names of these Archontic 'gods' tunes you into their frequency. That is the last thing you want when it allows for energetic synchronisation, attachment and perceptual influence. Initiates chant the names of their 'Gods' in their rituals for this very reason.

Vampires of the Woke

Paedophilia is another way that Archons absorb the energy of children. Paedophiles possessed by Archontic consciousness are used as the conduit during sexual abuse for discarnate Archons to vampire the energy of the young they desire so much. Stupendous numbers of children disappear every year never to be seen again although you would never know from the media. Imagine how much low-vibrational energy has been generated by children during the 'Covid' hoax when so many have become depressed and psychologically destroyed to the point of killing themselves. Shocking numbers of children are now taken by the state from loving parents to be handed to others. I can tell you from long experience of researching this since 1996 that many end up with paedophiles and assets of the Cult through corrupt and Cult-owned social services which in the reframing era has hired many psychopaths and emotionless automatons to do the job. Children are even stolen to order using spurious reasons to take them by the corrupt and secret (because they're corrupt) 'family courts'. I have written in detail in other books, starting with *The Biggest Secret* in 1997, about the ubiquitous connections between the political, corporate, government, intelligence and military elites (Cult operatives) and Satanism and paedophilia. If you go deep enough both networks have an interlocking leadership. The Woke mentality has been developed by the Cult for many reasons: To promote almost every aspect of its agenda; to hijack the traditional political left and turn it fascist; to divide and rule; and to target agenda pushbackers. But there are other reasons which relate to what I am describing here. How many happy and joyful Wokers do you ever see especially at the extreme end? They are a mental and psychological mess consumed by emotional stress and constantly emotionally cocked for the next explosion of indignation at someone referring to a female as a female. They are walking, talking, batteries as Morpheus might say emitting frequencies which both enslave them in low-vibrational bubbles of perceptual limitation and feed the Archons. Add to this the hatred claimed to be love; fascism claimed to 'anti-fascism', racism claimed to be 'anti-racism';

exclusion claimed to inclusion; and the abuse-filled Internet trolling. You have a purpose-built Archontic energy system with not a wind turbine in sight and all founded on Archontic *inversion*. We have whole generations now manipulated to serve the Archons with their actions and energy. They will be doing so their entire adult lives unless they snap out of their Archon-induced trance. Is it really a surprise that Cult billionaires and corporations put so much money their way? Where is the energy of joy and laughter, including laughing at yourself which is confirmation of your own emotional security? Mark Twain said: 'The human race has one really effective weapon, and that is laughter.' We must use it all the time. Woke has destroyed comedy because it has no humour, no joy, sense of irony, or self-deprecation. Its energy is dense and intense. *Mmmmm*, lunch says the Archontic frequency. Rudolf Steiner (1861-1925) was the Austrian philosopher and famous esoteric thinker who established Waldorf education or Steiner schools to treat children like unique expressions of consciousness and not minds to be programmed with the perceptions determined by authority. I'd been writing about this energy vampiring for decades when I was sent in 2016 a quote by Steiner. He was spot on:

> There are beings in the spiritual realms for whom anxiety and fear emanating from human beings offer welcome food. When humans have no anxiety and fear, then these creatures starve. If fear and anxiety radiates from people and they break out in panic, then these creatures find welcome nutrition and they become more and more powerful. These beings are hostile towards humanity. Everything that feeds on negative feelings, on anxiety, fear and superstition, despair or doubt, are in reality hostile forces in super-sensible worlds, launching cruel attacks on human beings, while they are being fed ... These are exactly the feelings that belong to contemporary culture and materialism; because it estranges people from the spiritual world, it is especially suited to evoke hopelessness and fear of the unknown in people, thereby calling up the above mentioned hostile forces against them.

Pause for a moment from this perspective and reflect on what has happened in the world since the start of 2020. Not only will pennies drop, but billion dollar bills. We see the same theme from Don Juan Matus, a Yaqui Indian shaman in Mexico and the information source for Peruvian-born writer, Carlos Castaneda, who wrote a series of

books from the 1960s to 1990s. Don Juan described the force manipulating human society and his name for the Archons was the predator:

> We have a predator that came from the depths of the cosmos and took over the rule of our lives. Human beings are its prisoners. The predator is our lord and master. It has rendered us docile, helpless. If we want to protest, it suppresses our protest. If we want to act independently, it demands that we don't do so ... indeed we are held prisoner!
>
> They took us over because we are food to them, and they squeeze us mercilessly because we are their sustenance. Just as we rear chickens in coops, the predators rear us in human coops, humaneros. Therefore, their food is always available to them.
>
> Different cultures, different eras, same recurring theme.

The 'ennoia' dilemma

Nag Hammadi Gnostic manuscripts say that Archon consciousness has no 'ennoia'. This is directly translated as 'intentionality', but I'll use the term 'creative imagination'. The *All That Is* in awareness of itself is the source of all creativity – all possibility – and the more disconnected you are from that source the more you are subsequently denied 'creative imagination'. Given that Archon consciousness is almost entirely disconnected it severely lacks creativity and has to rely on far more mechanical processes of thought and exploit the creative potential of those that do have 'ennoia'. You can see cases of this throughout human society. Archon consciousness almost entirely dominates the global banking system and if we study how that system works you will appreciate what I mean. Banks manifest 'money' out of nothing by issuing lines of 'credit' which is 'money' that has never, does not, and will never exist except in theory. It's a confidence trick. If you think 'credit' figures-on-a-screen 'money' is worth anything you accept it as payment. If you don't then the whole system collapses through lack of confidence in the value of that 'money'. Archontic bankers with no 'ennoia' are 'lending' 'money' that doesn't exist to humans that *do* have creativity – those that have the inspired ideas and create businesses and products. Archon banking feeds off human creativity

which it controls through 'money' creation and debt. Humans have the creativity and Archons exploit that for their own benefit and control while having none themselves. Archon Internet platforms like Facebook claim joint copyright of everything that creative users post and while Archontic minds like Zuckerberg may officially head that company it will be human creatives on the staff that provide the creative inspiration. When you have limitless 'money' you can then buy other companies established by creative humans. Witness the acquisition record of Facebook, Google and their like. Survey the Archon-controlled music industry and you see non-creative dark suit executives making their fortune from the human creativity of their artists. The cases are endless. Research the history of people like Gates and Zuckerberg and how their empires were built on exploiting the creativity of others. Archon minds cannot create out of nothing, but they are skilled (because they have to be) in what Gnostic texts call 'countermimicry'. They can imitate, but not innovate. Sabbatians trawl the creativity of others through backdoors they install in computer systems through their cybersecurity systems. Archon-controlled China is globally infamous for stealing intellectual property and I remember how Hong Kong, now part of China, became notorious for making counterfeit copies of the creativity of others – 'countermimicry'. With the now pervasive and all-seeing surveillance systems able to infiltrate any computer you can appreciate the potential for Archons to vampire the creativity of humans. Author John Lamb Lash wrote in his book about the Nag Hammadi texts, *Not In His Image*:

> Although they cannot originate anything, because they lack the divine factor of ennoia (intentionality), Archons can imitate with a vengeance. Their expertise is simulation (HAL, virtual reality). The Demiurge [Yaldabaoth] fashions a heaven world copied from the fractal patterns [of the original] ... His construction is celestial kitsch, like the fake Italianate villa of a Mafia don complete with militant angels to guard every portal.

This brings us to something that I have been speaking about since the turn of the millennium. Our reality is a simulation; a virtual reality that we think is real. No, I'm not kidding.

Human reality? Well, virtually

I had pondered for years about whether our reality is 'real' or some kind of construct. I remembered being immensely affected on a visit as a small child in the late 1950s to the then newly-opened Planetarium on the Marylebone Road in London which is now closed and part of the adjacent Madame Tussauds wax museum. It was in the middle of the day, but when the lights went out there was the night sky projected in the Planetarium's domed ceiling and it appeared to be so real. The experience never left me and I didn't know why until around the turn of the millennium when I became certain that our 'night sky' and entire reality is a projection, a virtual reality, akin to the illusory world portrayed in the *Matrix* movies. I looked at the sky one day in this period and it appeared to me like the domed roof of the Planetarium. The release of the first *Matrix* movie in 1999 also provided a synchronistic and perfect visual representation of where my mind had been going for a long time. I hadn't come across the Gnostic Nag Hammadi texts then. When I did years later the correlation was once again astounding. As I read Gnostic accounts from 1,600 years and more earlier it was clear that they were describing the same simulation phenomenon. They tell how the Yaldabaoth 'Demiurge' and Archons created a 'bad copy' of original reality to rule over all that were captured by its illusions and the body was a prison to trap consciousness in the 'bad copy' fake reality. Read how Gnostics describe the 'bad copy' and update that to current times and they are referring to what we would call today a virtual reality simulation.

Author John Lamb Lash said 'the Demiurge fashions a heaven world copied from the fractal patterns' of the original through expertise in 'HAL' or virtual reality simulation. Fractal patterns are part of the energetic information construct of our reality, a sort of blueprint. If these patterns were copied in computer terms it would indeed give you a copy of a 'natural' reality in a non-natural frequency and digital form. The principle is the same as making a copy of a website. The original website still exists, but now you can change the copy version to make it whatever you like and it can

become very different to the original website. Archons have done this with our reality, a *synthetic* copy of prime reality that still exists beyond the frequency walls of the simulation. Trapped within the illusions of this synthetic Matrix, however, were and are human consciousness and other expressions of prime reality and this is why the Archons via the Cult are seeking to make the human body synthetic and give us synthetic AI minds to complete the job of turning the entire reality synthetic including what we perceive to be the natural world. To quote Kurzweil: 'Nanobots will infuse all the matter around us with information. Rocks, trees, everything will become these intelligent creatures.' Yes, *synthetic* 'creatures' just as 'Covid' and other genetically-manipulating 'vaccines' are designed to make the human body synthetic. From this perspective it is obvious why Archons and their Cult are so desperate to infuse synthetic material into every human with their 'Covid' scam.

Let there be (electromagnetic) light

Yaldabaoth, the force that created the simulation, or Matrix, makes sense of the Gnostic reference to 'The Great Architect' and its use by Cult Freemasonry as the name of its deity. The designer of the Matrix in the movies is called 'The Architect' and that trilogy is jam-packed with symbolism relating to these subjects. I have contended for years that the angry Old Testament God (Yaldabaoth) is the 'God' being symbolically 'quoted' in the opening of Genesis as 'creating the world'. This is not the creation of prime reality – it's the creation of the *simulation*. The Genesis 'God' says: 'Let there be Light: and there was light.' But what is this 'Light'? I have said for decades that the speed of light (186,000 miles per second) is not the fastest speed possible as claimed by mainstream science and is in fact the frequency walls or outer limits of the Matrix. You can't have a fastest or slowest anything within all possibility when everything is possible. The human body is encoded to operate within the speed of light or *within the simulation* and thus we see only the tiny frequency band of visible *light*. Near-death experiencers who perceive reality outside the body during temporary 'death' describe a very different

form of light and this is supported by the Nag Hammadi texts. Prime reality beyond the simulation ('Upper Aeons' to the Gnostics) is described as a realm of incredible beauty, bliss, love and harmony – a realm of 'watery light' that is so powerful 'there are no shadows'. Our false reality of Archon control, which Gnostics call the 'Lower Aeons', is depicted as a realm with a different kind of 'light' and described in terms of chaos, 'Hell', 'the Abyss' and 'Outer Darkness', where trapped souls are tormented and manipulated by demons (relate that to the 'Covid' hoax alone). The watery light theme can be found in near-death accounts and it is not the same as *simulation* 'light' which is electromagnetic or radiation light within the speed of light – the 'Lower Aeons'. Simulation 'light' is the 'luminous fire' associated by Gnostics with the Archons. The Bible refers to Yaldabaoth as 'that old serpent, called the Devil, and Satan, which deceiveth the whole world' (Revelation 12:9). I think that making a simulated copy of prime reality ('countermimicry') and changing it dramatically while all the time manipulating humanity to believe it to be real could probably meet the criteria of deceiving the whole world. Then we come to the Cult god Lucifer – the *Light Bringer*. Lucifer is symbolic of Yaldabaoth, the bringer of radiation light that forms the bad copy simulation within the speed of light. 'He' is symbolised by the lighted torch held by the Statue of Liberty and in the name 'Illuminati'. Sabbatian-Frankism declares that Lucifer is the true god and Lucifer is the real god of Freemasonry honoured as their 'Great or Grand Architect of the Universe' (simulation).

I would emphasise, too, the way Archontic technologically-generated luminous fire of radiation has deluged our environment since I was a kid in the 1950s and changed the nature of The Field with which we constantly interact. Through that interaction technological radiation is changing us. The Smart Grid is designed to operate with immense levels of communication power with 5G expanding across the world and 6G, 7G, in the process of development. Radiation is the simulation and the Archontic manipulation system. Why wouldn't the Archon Cult wish to unleash radiation upon us to an ever-greater extreme to form

Kurzweil's 'cloud'? The plan for a synthetic human is related to the need to cope with levels of radiation beyond even anything we've seen so far. Biological humans would not survive the scale of radiation they have in their script. The Smart Grid is a technological sub-reality within the technological simulation to further disconnect five-sense perception from expanded consciousness. It's a technological prison of the mind.

Infusing the 'spirit of darkness'

A recurring theme in religion and native cultures is the manipulation of human genetics by a non-human force and most famously recorded as the biblical 'sons of god' (the god*s* plural in the original) who interbred with the daughters of men. The Nag Hammadi *Apocryphon of John* tells the same story this way:

> He [Yaldabaoth] sent his angels [Archons/demons] to the daughters of men, that they might take some of them for themselves and raise offspring for their enjoyment. And at first they did not succeed. When they had no success, they gathered together again and they made a plan together ... And the angels changed themselves in their likeness into the likeness of their mates, filling them with the spirit of darkness, which they had mixed for them, and with evil ... And they took women and begot children out of the darkness according to the likeness of their spirit.

Possession when a discarnate entity takes over a human body is an age-old theme and continues today. It's very real and I've seen it. Satanic and secret society rituals can create an energetic environment in which entities can attach to initiates and I've heard many stories of how people have changed their personality after being initiated even into lower levels of the Freemasons. I have been inside three Freemasonic temples, one at a public open day and two by just walking in when there was no one around to stop me. They were in Ryde, the town where I live, Birmingham, England, when I was with a group, and Boston, Massachusetts. They all felt the same energetically – dark, dense, low-vibrational and sinister. Demonic attachment can happen while the initiate has no idea what is going on. To them it's just a ritual to get in the Masons and do a bit of good

business. In the far more extreme rituals of Satanism human possession is even more powerful and they are designed to make possession possible. The hierarchy of the Cult is dictated by the power and perceived status of the possessing Archon. In this way the Archon hierarchy becomes the Cult hierarchy. Once the entity has attached it can influence perception and behaviour and if it attaches to the extreme then so much of its energy (information) infuses into the body information field that the hologram starts to reflect the nature of the possessing entity. This is the *Exorcist* movie type of possession when facial features change and it's known as shapeshifting. Islam's Jinn are said to be invisible tricksters who change shape, 'whisper', confuse and take human form. These are all traits of the Archons and other versions of the same phenomenon. Extreme possession could certainty infuse the 'spirit of darkness' into a partner during sex as the Nag Hammadi texts appear to describe. Such an infusion can change genetics which is also energetic information. Human genetics is information and the 'spirit of darkness' is information. Mix one with the other and change must happen. Islam has the concept of a 'Jinn baby' through possession of the mother and by Jinn taking human form. There are many ways that human genetics can be changed and remember that Archons have been aware all along of advanced techniques to do this. What is being done in human society today – and far more – was known about by Archons at the time of the 'fallen ones' and their other versions described in religions and cultures.

Archons and their human-world Cult are obsessed with genetics as we see today and they know this dictates how information is processed into perceived reality during a human life. They needed to produce a human form that would decode the simulation and this is symbolically known as 'Adam and Eve' who left the 'garden' (prime reality) and 'fell' into Matrix reality. The simulation is not a 'physical' construct (there is no 'physical'); it is a source of information. Think Wi-Fi again. The simulation is an energetic field encoded with information and body-brain systems are designed to decode that information encoded in wave or frequency form which

is transmitted to the brain as electrical signals. These are decoded by the brain to construct our sense of reality – an illusory 'physical' world that only exists in the brain or the mind. Virtual reality games mimic this process using the same sensory decoding system. Information is fed to the senses to decode a virtual reality that can appear so real, but isn't (Figs 18 and 19). Some scientists believe – and I agree with them – that what we perceive as 'physical' reality only exists when we are looking or observing. The act of perception or focus triggers the decoding systems which turn waveform information into holographic reality. When we are not observing something our reality reverts from a holographic state to a waveform state. This relates to the same principle as a falling tree not making a noise unless someone is there to hear it or decode it. The concept makes sense from the simulation perspective. A computer is not decoding all the information in a Wi-Fi field all the time and only decodes or brings into reality on the screen that part of Wi-Fi that it's decoding – focusing upon – at that moment.

Figure 18: Virtual reality technology 'hacks' into the body's five-sense decoding system.

Figure 19: The result can be experienced as very 'real'.

Interestingly, Professor Donald Hoffman at the Department of Cognitive Sciences at the University of California, Irvine, says that our experienced reality is like a computer interface that shows us only the level with which we interact while hiding all that exists beyond it: 'Evolution shaped us with a user interface that hides the truth. Nothing that we see is the truth – the very language of space and time and objects is the wrong language to describe reality.' He is correct in what he says on so many levels. Space and time are not a universal reality. They are a phenomenon of decoded *simulation* reality as part of the process of enslaving our sense of reality. Near-death experiencers report again and again how space and time did not exist as we perceive them once they were free of the body – body decoding systems. You can appreciate from this why Archons and their Cult are so desperate to entrap human attention in the five senses where we are in the Matrix and of the Matrix. Opening your mind to expanded states of awareness takes you beyond the information confines of the simulation and you become aware of knowledge and insights denied to you before. This is what we call 'awakening' – *awakening from the Matrix* – and in the final chapter I will relate this to current events.

Where are the 'aliens'?

A simulation would explain the so-called 'Fermi Paradox' named after Italian physicist Enrico Fermi (1901-1954) who created the first nuclear reactor. He considered the question of why there is such a lack of extraterrestrial activity when there are so many stars and planets in an apparently vast universe; but what if the night sky that we see, or think we do, is a simulated projection as I say? If you control the simulation and your aim is to hold humanity fast in essential ignorance would you want other forms of life including advanced life coming and going sharing information with humanity? Or would you want them to believe they were isolated and apparently alone? Themes of human isolation and apartness are common whether they be the perception of a lifeless universe or the fascist isolation laws of the 'Covid' era. Paradoxically the very

existence of a simulation means that we are not alone when some force had to construct it. My view is that experiences that people have reported all over the world for centuries with Reptilians and Grey entities are Archon phenomena as Nag Hammadi texts describe; and that benevolent 'alien' interactions are non-human groups that come in and out of the simulation by overcoming Archon attempts to keep them out. It should be highlighted, too, that Reptilians and Greys are obsessed with *genetics* and *technology* as related by cultural accounts and those who say they have been abducted by them. Technology is their way of overcoming some of the limitations in their creative potential and our technology-driven and controlled human society of today is *arch*etypical Archon-Reptilian-Grey modus operandi. Technocracy is really *Archon*tocracy. The Universe does not have to be as big as it appears with a simulation. There is no space or distance only information decoded into holographic reality. What we call 'space' is only the absence of holographic 'objects' and that 'space' is The Field of energetic information which connects everything into a single whole. The same applies with the artificially-generated information field of the simulation. The Universe is not big or small as a physical reality. It is decoded information, that's all, and its perceived size is decided by the way the simulation is encoded to make it appear. The entire night sky as we perceive it only exists in our brain and so where are those 'millions of light years'? The 'stars' on the ceiling of the Planetarium looked a vast distance away.

There's another point to mention about 'aliens'. I have been highlighting since the 1990s the plan to stage a fake 'alien invasion' to justify the centralisation of global power and a world military. Nazi scientist Werner von Braun, who was taken to America by Operation Paperclip after World War Two to help found NASA, told his American assistant Dr Carol Rosin about the Cult agenda when he knew he was dying in 1977. Rosin said that he told her about a sequence that would lead to total human control by a one-world government. This included threats from terrorism, rogue nations, meteors and asteroids before finally an 'alien invasion'. All of these

things, von Braun said, would be bogus and what I would refer to as a No-Problem-Reaction-Solution. Keep this in mind when 'the aliens are coming' is the new mantra. The aliens are not coming – they are *already here* and they have infiltrated human society while looking human. French-Canadian investigative journalist Serge Monast said in 1994 that he had uncovered a NASA/military operation called Project Blue Beam which fits with what Werner von Braun predicted. Monast died of a 'heart attack' in 1996 the day after he was arrested and spent a night in prison. He was 51. He said Blue Beam was a plan to stage an alien invasion that would include religious figures beamed holographically into the sky as part of a global manipulation to usher in a 'new age' of worshipping what I would say is the Cult 'god' Yaldabaoth in a one-world religion. Fake holographic asteroids are also said to be part of the plan which again syncs with von Braun. How could you stage an illusory threat from asteroids unless they were holographic inserts? This is pretty straightforward given the advanced technology outside the public arena and the fact that our 'physical' reality is holographic anyway. Information fields would be projected and we would decode them into the illusion of a 'physical' asteroid. If they can sell a global 'pandemic' with a 'virus' that doesn't exist what will humans not believe if government and media tell them?

All this is particularly relevant as I write with the Pentagon planning to release in June, 2021, information about 'UFO sightings'. I have been following the UFO story since the early 1990s and the common theme throughout has been government and military denials and cover up. More recently, however, the Pentagon has suddenly become more talkative and apparently open with Air Force pilot radar images released of unexplained craft moving and changing direction at speeds well beyond anything believed possible with human technology. Then, in March, 2021, former Director of National Intelligence John Ratcliffe said a Pentagon report months later in June would reveal a great deal of information about UFO sightings unknown to the public. He said the report would have 'massive implications'. The order to do this was included bizarrely

in a $2.3 trillion 'coronavirus' relief and government funding bill passed by the Trump administration at the end of 2020. I would add some serious notes of caution here. I have been pointing out since the 1990s that the US military and intelligence networks have long had craft – 'flying saucers' or anti-gravity craft – which any observer would take to be extraterrestrial in origin. Keeping this knowledge from the public allows craft flown by *humans* to be perceived as alien visitations. I am not saying that 'aliens' do not exist. I would be the last one to say that, but we have to be streetwise here. President Ronald Reagan told the UN General Assembly in 1987: 'I occasionally think how quickly our differences worldwide would vanish if we were facing an alien threat from outside this world.' That's the idea. Unite against a common 'enemy' with a common purpose behind your 'saviour force' (the Cult) as this age-old technique of mass manipulation goes global.

Science moves this way …

I could find only one other person who was discussing the simulation hypothesis publicly when I concluded it was real. This was Nick Bostrom, a Swedish-born philosopher at the University of Oxford, who has explored for many years the possibility that human reality is a computer simulation although his version and mine are not the same. Today the simulation and holographic reality hypothesis have increasingly entered the scientific mainstream. Well, the more open-minded mainstream, that is. Here are a few of the ever-gathering examples. American nuclear physicist Silas Beane led a team of physicists at the University of Bonn in Germany pursuing the question of whether we live in a simulation. They concluded that we probably do and it was likely based on a lattice of cubes. They found that cosmic rays align with that specific pattern. The team highlighted the Greisen–Zatsepin–Kuzmin (GZK) limit which refers to cosmic ray particle interaction with cosmic background radiation that creates an apparent boundary for cosmic ray particles. They say in a paper entitled 'Constraints on the Universe as a Numerical Simulation' that this 'pattern of constraint' is exactly what you

would find with a computer simulation. They also made the point that a simulation would create its own 'laws of physics' that would limit possibility. I've been making the same point for decades that the *perceived* laws of physics relate only to this reality, or what I would later call the simulation. When designers write codes to create computer and virtual reality games they are the equivalent of the laws of physics for that game. Players interact within the limitations laid out by the coding. In the same way those who wrote the codes for the simulation decided the laws of physics that would apply. These can be overridden by expanded states of consciousness, but not by those enslaved in only five-sense awareness where simulation codes rule. Overriding the codes is what people call 'miracles'. They are not. They are bypassing the encoded limits of the simulation. A population caught in simulation perception would have no idea that this was their plight. As the Bonn paper said: 'Like a prisoner in a pitch-black cell we would not be able to see the "walls" of our prison,' That's true if people remain mesmerised by the five senses. Open to expanded awareness and those walls become very clear. The main one is the speed of light.

American theoretical physicist James Gates is another who has explored the simulation question and found considerable evidence to support the idea. Gates was Professor of Physics at the University of Maryland, Director of The Center for String and Particle Theory, and on Barack Obama's Council of Advisors on Science and Technology. He and his team found *computer codes* of digital data embedded in the fabric of our reality. They relate to on-off electrical charges of 1 and 0 in the binary system used by computers. 'We have no idea what they are doing there', Gates said. They found within the energetic fabric mathematical sequences known as error-correcting codes or block codes that 'reboot' data to its original state or 'default settings' when something knocks it out of sync. Gates was asked if he had found a set of equations embedded in our reality indistinguishable from those that drive search engines and browsers and he said: 'That is correct.' Rich Terrile, director of the Centre for Evolutionary Computation and Automated Design at NASA's Jet

Propulsion Laboratory, has said publicly that he believes the Universe is a digital hologram that must have been created by a form of intelligence. I agree with that in every way. Waveform information is delivered electrically by the senses to the brain which constructs a *digital* holographic reality that we call the 'world'. This digital level of reality can be read by the esoteric art of numerology. Digital holograms are at the cutting edge of holographics today. We have digital technology everywhere designed to access and manipulate our digital level of perceived reality. Synthetic mRNA in 'Covid vaccines' has a digital component to manipulate the body's digital 'operating system'.

Reality is numbers

How many know that our reality can be broken down to numbers and codes that are the same as computer games? Max Tegmark, a physicist at the Massachusetts Institute of Technology (MIT), is the author of *Our Mathematical Universe* in which he lays out how reality can be entirely described by numbers and maths in the way that a video game is encoded with the 'physics' of computer games. Our world and computer virtual reality are essentially the same. Tegmark imagines the perceptions of characters in an advanced computer game when the graphics are so good they don't know they are in a game. They think they can bump into real objects (electromagnetic resistance in our reality), fall in love and feel emotions like excitement. When they began to study the apparently 'physical world' of the video game they would realise that everything was made of pixels (which have been found in our energetic reality as must be the case when on one level our world is digital). What computer game characters thought was physical 'stuff', Tegmark said, could actually be broken down into numbers:

> And we're exactly in this situation in our world. We look around and it doesn't seem that mathematical at all, but everything we see is made out of elementary particles like quarks and electrons. And what properties does an electron have? Does it have a smell or a colour or a texture? No! ... We physicists have come up with geeky names for [Electron] properties, like

electric charge, or spin, or lepton number, but the electron doesn't care what we call it, the properties are just numbers.

This is the illusory reality Gnostics were describing. This is the simulation. The A, C, G, and T codes of DNA have a binary value – A and C = 0 while G and T = 1. This has to be when the simulation is digital and the body must be digital to interact with it. Recurring mathematical sequences are encoded throughout reality and the body. They include the Fibonacci sequence in which the two previous numbers are added to get the next one, as in ... 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, etc. The sequence is encoded in the human face and body, proportions of animals, DNA, seed heads, pine cones, trees, shells, spiral galaxies, hurricanes and the number of petals in a flower. The list goes on and on. There are fractal patterns – a 'never-ending pattern that is infinitely complex and self-similar across all scales in the as above, so below, principle of holograms. These and other famous recurring geometrical and mathematical sequences such as Phi, Pi, Golden Mean, Golden Ratio and Golden Section are *computer codes* of the simulation. I had to laugh and give my head a shake the day I finished this book and it went into the production stage. I was sent an article in *Scientific American* published in April, 2021, with the headline 'Confirmed! We Live in a Simulation'. Two decades after I first said our reality is a simulation and the speed of light is it's outer limit the article suggested that we do live in a simulation and that the speed of light is its outer limit. I left school at 15 and never passed a major exam in my life while the writer was up to his eyes in qualifications. As I will explain in the final chapter *knowing* is far better than thinking and they come from very different sources. The article rightly connected the speed of light to the processing speed of the 'Matrix' and said what has been in my books all this time … 'If we are in a simulation, as it appears, then space is an abstract property written in code. It is not real'. No it's not and if we live in a simulation something created it and it wasn't *us*. 'That David Icke says we are manipulated by aliens' – he's crackers.'

Wow …

The reality that humanity thinks is so real is an illusion. Politicians, governments, scientists, doctors, academics, law enforcement, media, school and university curriculums, on and on, are all founded on a world that *does not exist* except as a simulated prison cell. Is it such a stretch to accept that 'Covid' doesn't exist when our entire 'physical' reality doesn't exist? Revealed here is the knowledge kept under raps in the Cult networks of compartmentalised secrecy to control humanity's sense of reality by inducing the population to believe in a reality that's not real. If it wasn't so tragic in its experiential consequences the whole thing would be hysterically funny. None of this is new to Renegade Minds. Ancient Greek philosopher Plato (about 428 to about 347BC) was a major influence on Gnostic belief and he described the human plight thousands of years ago with his Allegory of the Cave. He told the symbolic story of prisoners living in a cave who had never been outside. They were chained and could only see one wall of the cave while behind them was a fire that they could not see. Figures walked past the fire casting shadows on the prisoners' wall and those moving shadows became their sense of reality. Some prisoners began to study the shadows and were considered experts on them (today's academics and scientists), but what they studied was only an illusion (today's academics and scientists). A prisoner escaped from the cave and saw reality as it really is. When he returned to report this revelation they didn't believe him, called him mad and threatened to kill him if he tried to set them free. Plato's tale is not only a brilliant analogy of the human plight and our illusory reality. It describes, too, the dynamics of the 'Covid' hoax. I have only skimmed the surface of these subjects here. The aim of this book is to crisply connect all essential dots to put what is happening today into its true context. All subject areas and their connections in this chapter are covered in great evidential detail in *Everything You Need To Know, But Have Never Been Told* and *The Answer*.

They say that bewildered people 'can't see the forest for the trees'. Humanity, however, can't see the forest for the *twigs*. The five senses

see only twigs while Renegade Minds can see the forest and it's the forest where the answers lie with the connections that reveals. Breaking free of perceptual programming so the forest can be seen is the way we turn all this around. Not breaking free is how humanity got into this mess. The situation may seem hopeless, but I promise you it's not. We are a perceptual heartbeat from paradise if only we knew.

CHAPTER TWELVE

Escaping Wetiko

Life is simply a vacation from the infinite
Dean Cavanagh

Renegade Minds weave the web of life and events and see common themes in the apparently random. They are always there if you look for them and their pursuit is aided by incredible synchronicity that comes when your mind is open rather than mesmerised by what it thinks it can see.

Infinite awareness is infinite possibility and the more of infinite possibility that we access the more becomes infinitely possible. That may be stating the apparently obvious, but it is a devastatingly-powerful fact that can set us free. We are a point of attention within an infinity of consciousness. The question is how much of that infinity do we choose to access? How much knowledge, insight, awareness, wisdom, do we want to connect with and explore? If your focus is only in the five senses you will be influenced by a fraction of infinite awareness. I mean a range so tiny that it gives new meaning to infinitesimal. Limitation of self-identity and a sense of the possible limit accordingly your range of consciousness. We are what we think we are. Life is what we think it is. The dream is the dreamer and the dreamer is the dream. Buddhist philosophy puts it this way: 'As a thing is viewed, so it appears.' Most humans live in the realm of touch, taste, see, hear, and smell and that's the limit of their sense of the possible and sense of self. Many will follow a religion and speak of a God in his heaven, but their lives are still

dominated by the five senses in their perceptions and actions. The five senses become the arbiter of everything. When that happens all except a smear of infinity is sealed away from influence by the rigid, unyielding, reality bubbles that are the five-sense human or Phantom Self. Archon Cult methodology is to isolate consciousness within five-sense reality – the simulation – and then program that consciousness with a sense of self and the world through a deluge of life-long information designed to instil the desired perception that allows global control. Efforts to do this have increased dramatically with identity politics as identity bubbles are squeezed into the minutiae of five-sense detail which disconnect people even more profoundly from the infinite 'I'.

Five-sense focus and self-identity are like a firewall that limits access to the infinite realms. You only perceive one radio or television station and no other. We'll take that literally for a moment. Imagine a vast array of stations giving different information and angles on reality, but you only ever listen to one. Here we have the human plight in which the population is overwhelmingly confined to CultFM. This relates only to the frequency range of CultFM and limits perception and insight to that band – limits *possibility* to that band. It means you are connecting with an almost imperceptibly minuscule range of possibility and creative potential within the infinite Field. It's a world where everything seems apart from everything else and where synchronicity is rare. Synchronicity is defined in the dictionary as 'the happening by chance of two or more related or similar events at the same time'. Use of 'by chance' betrays a complete misunderstanding of reality. Synchronicity is not 'by chance'. As people open their minds, or 'awaken' to use the term, they notice more and more coincidences in their lives, bits of 'luck', apparently miraculous happenings that put them in the right place at the right time with the right people. Days become peppered with 'fancy meeting you here' and 'what are the chances of that?' My entire life has been lived like this and ever more so since my own colossal awakening in 1990 and 91 which transformed my sense of reality. Synchronicity is not 'by chance'; it is by accessing expanded

realms of possibility which allow expanded potential for manifestation. People broadcasting the same vibe from the same openness of mind tend to be drawn 'by chance' to each other through what I call frequency magnetism and it's not only people. In the last more than 30 years incredible synchronicity has also led me through the Cult maze to information in so many forms and to crucial personal experiences. These 'coincidences' have allowed me to put the puzzle pieces together across an enormous array of subjects and situations. Those who have breached the bubble of five-sense reality will know exactly what I mean and this escape from the perceptual prison cell is open to everyone whenever they make that choice. This may appear super-human when compared with the limitations of 'human', but it's really our natural state. 'Human' as currently experienced is consciousness in an unnatural state of induced separation from the infinity of the whole. I'll come to how this transformation into unity can be made when I have described in more detail the force that holds humanity in servitude by denying this access to infinite self.

The Wetiko factor

I have been talking and writing for decades about the way five-sense mind is systematically barricaded from expanded awareness. I have used the analogy of a computer (five-sense mind) and someone at the keyboard (expanded awareness). Interaction between the computer and the operator is symbolic of the interaction between five-sense mind and expanded awareness. The computer directly experiences the Internet and the operator experiences the Internet via the computer which is how it's supposed to be – the two working as one. Archons seek to control that point where the operator connects with the computer to stop that interaction (Fig 20). Now the operator is banging the keyboard and clicking the mouse, but the computer is not responding and this happens when the computer is taken over – *possessed* – by an appropriately-named computer 'virus'. The operator has lost all influence over the computer which goes its own way making decisions under the control of the 'virus'. I have

just described the dynamic through which the force known to Gnostics as Yaldabaoth and Archons disconnects five-sense mind from expanded awareness to imprison humanity in perceptual servitude.

Figure 20: The mind 'virus' I have been writing about for decades seeks to isolate five-sense mind (the computer) from the true 'I'. (Image by Neil Hague).

About a year ago I came across a Native American concept of Wetiko which describes precisely the same phenomenon. Wetiko is the spelling used by the Cree and there are other versions including wintiko and windigo used by other tribal groups. They spell the name with lower case, but I see Wetiko as a proper noun as with Archons and prefer a capital. I first saw an article about Wetiko by writer and researcher Paul Levy which so synced with what I had been writing about the computer/operator disconnection and later the Archons. I then read his book, the fascinating *Dispelling Wetiko, Breaking the Spell of Evil*. The parallels between what I had concluded long before and the Native American concept of Wetiko were so clear and obvious that it was almost funny. For Wetiko see the Gnostic Archons for sure and the Jinn, the Predators, and every other name for a force of evil, inversion and chaos. Wetiko is the Native American name for the force that divides the computer from

the operator (Fig 21). Indigenous author Jack D. Forbes, a founder of the Native American movement in the 1960s, wrote another book about Wetiko entitled *Columbus And Other Cannibals – The Wetiko Disease of Exploitation, Imperialism, and Terrorism* which I also read. Forbes says that Wetiko refers to an evil person or spirit 'who terrorizes other creatures by means of terrible acts, including cannibalism'. Zulu shaman Credo Mutwa told me that African accounts tell how cannibalism was brought into the world by the Chitauri 'gods' – another manifestation of Wetiko. The distinction between 'evil person or spirit' relates to Archons/Wetiko possessing a human or acting as pure consciousness. Wetiko is said to be a sickness of the soul or spirit and a state of being that takes but gives nothing back – the Cult and its operatives perfectly described. Black Hawk, a Native American war leader defending their lands from confiscation, said European invaders had 'poisoned hearts' – Wetiko hearts – and that this would spread to native societies. Mention of the heart is very significant as we shall shortly see. Forbes writes: 'Tragically, the history of the world for the past 2,000 years is, in great part, the story of the epidemiology of the wetiko disease.' Yes, and much longer. Forbes is correct when he says: 'The wetikos destroyed Egypt and Babylon and Athens and Rome and Tenochtitlan [capital of the Aztec empire] and perhaps now they will destroy the entire earth.' Evil, he said, is the number one export of a Wetiko culture – see its globalisation with 'Covid'. Constant war, mass murder, suffering of all kinds, child abuse, Satanism, torture and human sacrifice are all expressions of Wetiko and the Wetiko possessed. The world is Wetiko made manifest, *but it doesn't have to be*. There is a way out of this even now.

Figure 21: The mind 'virus' is known to Native Americans as 'Wetiko'. (Image by Neil Hague).

Cult of Wetiko

Wetiko is the Yaldabaoth frequency distortion that seeks to attach to human consciousness and absorb it into its own. Once this connection is made Wetiko can drive the perceptions of the target which they believe to be coming from their own mind. All the horrors of history and today from mass killers to Satanists, paedophiles like Jeffrey Epstein and other psychopaths, are the embodiment of Wetiko and express its state of being in all its grotesqueness. The Cult is Wetiko incarnate, Yaldabaoth incarnate, and it seeks to facilitate Wetiko assimilation of humanity in totality into its distortion by manipulating the population into low frequency states that match its own. Paul Levy writes: 'Holographically enforced within the psyche of every human being the wetiko virus pervades and underlies the entire field of consciousness, and can therefore potentially manifest through any one of us at any moment if we are not mindful.' The 'Covid' hoax has achieved this with many people, but others have not fallen into Wetiko's frequency lair. Players in the 'Covid' human catastrophe including Gates, Schwab, Tedros, Fauci, Whitty, Vallance, Johnson, Hancock, Ferguson, Drosten, and all the rest, including the psychopath psychologists, are expressions of Wetiko. This is why

they have no compassion or empathy and no emotional consequence for what they do that would make them stop doing it. Observe all the people who support the psychopaths in authority against the Pushbackers despite the damaging impact the psychopaths have on their own lives and their family's lives. You are again looking at Wetiko possession which prevents them seeing through the lies to the obvious scam going on. *Why can't they see it?* Wetiko won't let them see it. The perceptual divide that has now become a chasm is between the Wetikoed and the non-Wetikoed.

Paul Levy describes Wetiko in the same way that I have long described the Archontic force. They are the same distorted consciousness operating across dimensions of reality: '… the subtle body of wetiko is not located in the third dimension of space and time, literally existing in another dimension … it is able to affect ordinary lives by mysteriously interpenetrating into our three-dimensional world.' Wetiko does this through its incarnate representatives in the Cult and by weaving itself into The Field which on our level of reality is the electromagnetic information field of the simulation or Matrix. More than that, the simulation *is* Wetiko / Yaldabaoth. Caleb Scharf, Director of Astrobiology at Columbia University, has speculated that 'alien life' could be so advanced that it has transcribed itself into the quantum realm to become what we call physics. He said intelligence indistinguishable from the fabric of the Universe would solve many of its greatest mysteries:

> Perhaps hyper-advanced life isn't just external. Perhaps it's already all around. It is embedded in what we perceive to be physics itself, from the root behaviour of particles and fields to the phenomena of complexity and emergence ... In other words, life might not just be in the equations. It might BE the equations [My emphasis].

Scharf said it is possible that 'we don't recognise advanced life because it forms an integral and unsuspicious part of what we've considered to be the natural world'. I agree. Wetiko/Yaldabaoth *is* the simulation. We are literally in the body of the beast. But that doesn't mean it has to control us. We all have the power to overcome Wetiko

influence and the Cult knows that. I doubt it sleeps too well because it knows that.

Which Field?

This, I suggest, is how it all works. There are two Fields. One is the fierce electromagnetic light of the Matrix within the speed of light; the other is the 'watery light' of The Field beyond the walls of the Matrix that connects with the Great Infinity. Five-sense mind and the decoding systems of the body attach us to the Field of Matrix light. They have to or we could not experience this reality. Five-sense mind sees only the Matrix Field of information while our expanded consciousness is part of the Infinity Field. When we open our minds, and most importantly our hearts, to the Infinity Field we have a mission control which gives us an expanded perspective, a road map, to understand the nature of the five-sense world. If we are isolated only in five-sense mind there is no mission control. We're on our own trying to understand a world that's constantly feeding us information to ensure we do not understand. People in this state can feel 'lost' and bewildered with no direction or radar. You can see ever more clearly those who are influenced by the Fields of Big Infinity or little five-sense mind simply by their views and behaviour with regard to the 'Covid' hoax. We have had this division throughout known human history with the mass of the people on one side and individuals who could see and intuit beyond the walls of the simulation – Plato's prisoner who broke out of the cave and saw reality for what it is. Such people have always been targeted by Wetiko/Archon-possessed authority, burned at the stake or demonised as mad, bad and dangerous. The Cult today and its global network of 'anti-hate', 'anti-fascist' Woke groups are all expressions of Wetiko attacking those exposing the conspiracy, 'Covid' lies and the 'vaccine' agenda.

Woke as a whole is Wetiko which explains its black and white mentality and how at one it is with the Wetiko-possessed Cult. Paul Levy said: 'To be in this paradigm is to still be under the thrall of a two-valued logic – where things are either true or false – of a

wetikoized mind.' Wetiko consciousness is in a permanent rage, therefore so is Woke, and then there is Woke inversion and contradiction. 'Anti-fascists' act like fascists because fascists *and* 'anti-fascists' are both Wetiko at work. Political parties act the same while claiming to be different for the same reason. Secret society and satanic rituals are attaching initiates to Wetiko and the cold, ruthless, psychopathic mentality that secures the positions of power all over the world is Wetiko. Reframing 'training programmes' have the same cumulative effect of attaching Wetiko and we have their graduates described as automatons and robots with a cold, psychopathic, uncaring demeanour. They are all traits of Wetiko possession and look how many times they have been described in this book and elsewhere with regard to personnel behind 'Covid' including the police and medical profession. Climbing the greasy pole in any profession in a Wetiko society requires traits of Wetiko to get there and that is particularly true of politics which is not about fair competition and pre-eminence of ideas. It is founded on how many backs you can stab and arses you can lick. This culminated in the global 'Covid' coordination between the Wetiko possessed who pulled it off in all the different countries without a trace of empathy and compassion for their impact on humans. Our sight sense can see only holographic form and not the Field which connects holographic form. Therefore we perceive 'physical' objects with 'space' in between. In fact that 'space' is energy/consciousness operating on multiple frequencies. One of them is Wetiko and that connects the Cult psychopaths, those who submit to the psychopaths, and those who serve the psychopaths in the media operations of the world. Wetiko is Gates. Wetiko is the mask-wearing submissive. Wetiko is the fake journalist and 'fact-checker'. The Wetiko Field is coordinating the whole thing. Psychopaths, gofers, media operatives, 'anti-hate' hate groups, 'fact-checkers' and submissive people work as one unit *even without human coordination* because they are attached to the *same* Field which is organising it all (Fig 22). Paul Levy is here describing how Wetiko-possessed people are drawn together and refuse to let any information breach their rigid

perceptions. He was writing long before 'Covid', but I think you will recognise followers of the 'Covid' religion *oh just a little bit*:

People who are channelling the vibratory frequency of wetiko align with each other through psychic resonance to reinforce their unspoken shared agreement so as to uphold their deranged view of reality. Once an unconscious content takes possession of certain individuals, it irresistibly draws them together by mutual attraction and knits them into groups tied together by their shared madness that can easily swell into an avalanche of insanity.

A psychic epidemic is a closed system, which is to say that it is insular and not open to any new information or informing influences from the outside world which contradict its fixed, limited, and limiting perspective.

There we have the Woke mind and the 'Covid' mind. Compatible resonance draws the awakening together, too, which is clearly happening today.

Figure 22: The Wetiko Field from which the Cult pyramid and its personnel are made manifest. (Image by Neil Hague).

Spiritual servitude

Wetiko doesn't care about humans. It's not human; it just possesses humans for its own ends and the effect (depending on the scale of

possession) can be anything from extreme psychopathy to unquestioning obedience. Wetiko's worst nightmare is for human consciousness to expand beyond the simulation. Everything is focussed on stopping that happening through control of information, thus perception, thus frequency. The 'education system', media, science, medicine, academia, are all geared to maintaining humanity in five-sense servitude as is the constant stimulation of low-vibrational mental and emotional states (see 'Covid'). Wetiko seeks to dominate those subconscious spaces between five-sense perception and expanded consciousness where the computer meets the operator. From these subconscious hiding places Wetiko speaks to us to trigger urges and desires that we take to be our own and manipulate us into anything from low-vibrational to psychopathic states. Remember how Islam describes the Jinn as invisible tricksters that 'whisper' and confuse. Wetiko is the origin of the 'trickster god' theme that you find in cultures all over the world. Jinn, like the Archons, are Wetiko which is terrified of humans awakening and reconnecting with our true self for then its energy source has gone. With that the feedback loop breaks between Wetiko and human perception that provides the energetic momentum on which its very existence depends as a force of evil. Humans are both its target and its source of survival, but only if we are operating in low-vibrational states of fear, hate, depression and the background anxiety that most people suffer. We are Wetiko's target because we are its key to survival. It needs us, not the other way round. Paul Levy writes:

A vampire has no intrinsic, independent, substantial existence in its own right; it only exists in relation to us. The pathogenic, vampiric mind-parasite called wetiko is nothing in itself – not being able to exist from its own side – yet it has a 'virtual reality' such that it can potentially destroy our species …

…The fact that a vampire is not reflected by a mirror can also mean that what we need to see is that there's nothing, no-thing to see, other than ourselves. The fact that wetiko is the expression of something inside of us means that the cure for wetiko is with us as well. The critical issue is finding this cure within us and then putting it into effect.

Evil begets evil because if evil does not constantly expand and find new sources of energetic sustenance its evil, its *distortion*, dies with the assimilation into balance and harmony. Love is the garlic to Wetiko's vampire. Evil, the absence of love, cannot exist in the presence of love. I think I see a way out of here. I have emphasised so many times over the decades that the Archons/Wetiko and their Cult are not all powerful. *They are not*. I don't care how it looks even now *they are not*. I have not called them little boys in short trousers for effect. I have said it because it is true. Wetiko's insatiable desire for power over others is not a sign of its omnipotence, but its insecurity. Paul Levy writes: 'Due to the primal fear which ultimately drives it and which it is driven to cultivate, wetiko's body politic has an intrinsic and insistent need for centralising power and control so as to create imagined safety for itself.' *Yeeeeeees!* Exactly! Why does Wetiko want humans in an ongoing state of fear? Wetiko itself *is* fear and it is petrified of love. As evil is an absence of love, so love is an absence of fear. Love conquers all and *especially* Wetiko which *is* fear. Wetiko brought fear into the world when it wasn't here before. *Fear* was the 'fall', the fall into low-frequency ignorance and illusion – fear is **F**alse **E**motion **A**ppearing **R**eal. The simulation is driven and energised by fear because Wetiko/Yaldabaoth (fear) *are* the simulation. Fear is the absence of love and Wetiko is the absence of love.

Wetiko today

We can now view current events from this level of perspective. The 'Covid' hoax has generated momentous amounts of ongoing fear, anxiety, depression and despair which have empowered Wetiko. No wonder people like Gates have been the instigators when they are Wetiko incarnate and exhibit every trait of Wetiko in the extreme. See how cold and unemotional these people are like Gates and his cronies, how dead of eye they are. That's Wetiko. Sabbatians are Wetiko and everything they control including the World Health Organization, Big Pharma and the 'vaccine' makers, national 'health'

hierarchies, corporate media, Silicon Valley, the banking system, and the United Nations with its planned transformation into world government. All are controlled and possessed by the Wetiko distortion into distorting human society in its image. We are with this knowledge at the gateway to understanding the world. Divisions of race, culture, creed and sexuality are diversions to hide the real division between those possessed and influenced by Wetiko and those that are not. The 'Covid' hoax has brought both clearly into view. Human behaviour is not about race. Tyrants and dictatorships come in all colours and creeds. What unites the US president bombing the innocent and an African tribe committing genocide against another as in Rwanda? What unites them? *Wetiko*. All wars are Wetiko, all genocide is Wetiko, all hunger over centuries in a world of plenty is Wetiko. Children going to bed hungry, including in the West, is Wetiko. Cult-generated Woke racial divisions that focus on the body are designed to obscure the reality that divisions in behaviour are manifestations of mind, not body. Obsession with body identity and group judgement is a means to divert attention from the real source of behaviour – mind and perception. Conflict sown by the Woke both within themselves and with their target groups are Wetiko providing lunch for itself through still more agents of the division, chaos, and fear on which it feeds. The Cult is seeking to assimilate the entirety of humanity and all children and young people into the Wetiko frequency by manipulating them into states of fear and despair. Witness all the suicide and psychological unravelling since the spring of 2020. Wetiko psychopaths want to impose a state of unquestioning obedience to authority which is no more than a conduit for Wetiko to enforce its will and assimilate humanity into itself. It needs us to believe that resistance is futile when it fears resistance and even more so the game-changing non-cooperation with its impositions. It can use violent resistance for its benefit. Violent impositions and violent resistance are *both* Wetiko. The Power of Love with its Power of No will sweep Wetiko from our world. Wetiko and its Cult know that. They just don't want us to know.

AI Wetiko

This brings me to AI or artificial intelligence and something else Wetikos don't want us to know. What is AI *really*? I know about computer code algorithms and AI that learns from data input. These, however, are more diversions, the expeditionary force, for the real AI that they want to connect to the human brain as promoted by Silicon Valley Wetikos like Kurzweil. What is this AI? It is the frequency of *Wetiko*, the frequency of the Archons. The connection of AI to the human brain is the connection of the Wetiko frequency to create a Wetiko hive mind and complete the job of assimilation. The hive mind is planned to be controlled from Israel and China which are both 100 percent owned by Wetiko Sabbatians. The assimilation process has been going on minute by minute in the 'smart' era which fused with the 'Covid' era. We are told that social media is scrambling the minds of the young and changing their personality. This is true, but what is social media? Look more deeply at how it works, how it creates divisions and conflict, the hostility and cruelty, the targeting of people until they are destroyed. That's Wetiko. Social media is manipulated to tune people to the Wetiko frequency with all the emotional exploitation tricks employed by platforms like Facebook and its Wetiko front man, Zuckerberg. Facebook's Instagram announced a new platform for children to overcome a legal bar on them using the main site. This is more Wetiko exploitation and manipulation of kids. Amnesty International likened the plan to foxes offering to guard the henhouse and said it was incompatible with human rights. Since when did Wetiko or Zuckerberg (I repeat myself) care about that? Would Brin and Page at Google, Wojcicki at YouTube, Bezos at Amazon and whoever the hell runs Twitter act as they do if they were not channelling Wetiko? Would those who are developing technologies for no other reason than human control? How about those designing and selling technologies to kill people and Big Pharma drug and 'vaccine' producers who know they will end or devastate lives? Quite a thought for these people to consider is that if you are Wetiko in a human life you are Wetiko on the 'other side' unless your frequency

changes and that can only change by a change of perception which becomes a change of behaviour. Where Gates is going does not bear thinking about although perhaps that's exactly where he wants to go. Either way, that's where he's going. His frequency will make it so.

The frequency lair

I have been saying for a long time that a big part of the addiction to smartphones and devices is that a frequency is coming off them that entraps the mind. People spend ages on their phones and sometimes even a minute or so after they put them down they pick them up again and it all repeats. 'Covid' lockdowns will have increased this addiction a million times for obvious reasons. Addictions to alcohol overindulgence and drugs are another way that Wetiko entraps consciousness to attach to its own. Both are symptoms of low-vibrational psychological distress which alcoholism and drug addiction further compound. Do we think it's really a coincidence that access to them is made so easy while potions that can take people into realms beyond the simulation are banned and illegal? I have explored smartphone addiction in other books, the scale is mind-blowing, and that level of addiction does not come without help. Tech companies that make these phones are Wetiko and they will have no qualms about destroying the minds of children. We are seeing again with these companies the Wetiko perceptual combination of psychopathic enforcers and weak and meek unquestioning compliance by the rank and file.

The global Smart Grid is the Wetiko Grid and it is crucial to complete the Cult endgame. The simulation is radiation and we are being deluged with technological radiation on a devastating scale. Wetiko frauds like Elon Musk serve Cult interests while occasionally criticising them to maintain his street-cred. 5G and other forms of Wi-Fi are being directed at the earth from space on a volume and scale that goes on increasing by the day. Elon Musk's (officially) SpaceX Starlink project is in the process of putting tens of thousands of satellites in low orbit to cover every inch of the planet with 5G and other Wi-Fi to create Kurzweil's global 'cloud' to which the

human mind is planned to be attached very soon. SpaceX has approval to operate 12,000 satellites with more than 1,300 launched at the time of writing and applications filed for *30,000* more. Other operators in the Wi-Fi, 5G, low-orbit satellite market include OneWeb (UK), Telesat (Canada), and AST & Science (US). Musk tells us that AI could be the end of humanity and then launches a company called Neuralink to connect the human brain to computers. Musk's (in theory) Tesla company is building electric cars and the driverless vehicles of the smart control grid. As frauds and bullshitters go Elon Musk in my opinion is Major League.

5G and technological radiation in general are destructive to human health, genetics and psychology and increasing the strength of artificial radiation underpins the five-sense perceptual bubbles which are themselves expressions of radiation or electromagnetism. Freedom activist John Whitehead was so right with his 'databit by databit, we are building our own electronic concentration camps'. The Smart Grid and 5G is a means to control the human mind and infuse perceptual information into The Field to influence anyone in sync with its frequency. You can change perception and behaviour en masse if you can manipulate the population into those levels of frequency and this is happening all around us today. The arrogance of Musk and his fellow Cult operatives knows no bounds in the way that we see with Gates. Musk's satellites are so many in number already they are changing the night sky when viewed from Earth. The astronomy community has complained about this and they have seen nothing yet. Some consequences of Musk's Wetiko hubris include: Radiation; visible pollution of the night sky; interference with astronomy and meteorology; ground and water pollution from intensive use of increasingly many spaceports; accumulating space debris; continual deorbiting and burning up of aging satellites, polluting the atmosphere with toxic dust and smoke; and ever-increasing likelihood of collisions. A collective public open letter of complaint to Musk said:

> We are writing to you … because SpaceX is in process of surrounding the Earth with a network of thousands of satellites whose very purpose is to irradiate every square inch of the

Earth. SpaceX, like everyone else, is treating the radiation as if it were not there. As if the mitochondria in our cells do not depend on electrons moving undisturbed from the food we digest to the oxygen we breathe.

As if our nervous systems and our hearts are not subject to radio frequency interference like any piece of electronic equipment. As if the cancer, diabetes, and heart disease that now afflict a majority of the Earth's population are not metabolic diseases that result from interference with our cellular machinery. As if insects everywhere, and the birds and animals that eat them, are not starving to death as a result.

People like Musk and Gates believe in their limitless Wetiko arrogance that they can do whatever they like to the world because they own it. Consequences for humanity are irrelevant. It's absolutely time that we stopped taking this shit from these self-styled masters of the Earth when you consider where this is going.

Why is the Cult so anti-human?

I hear this question often: Why would they do this when it will affect them, too? Ah, but will it? Who is this *them*? Forget their bodies. They are just vehicles for Wetiko consciousness. When you break it all down to the foundations we are looking at a state of severely distorted consciousness targeting another state of consciousness for assimilation. The rest is detail. The simulation is the fly-trap in which unique sensations of the five senses create a cycle of addiction called reincarnation. Renegade Minds see that everything which happens in our reality is a smaller version of the whole picture in line with the holographic principle. Addiction to the radiation of smart technology is a smaller version of addiction to the whole simulation. Connecting the body/brain to AI is taking that addiction on a giant step further to total ongoing control by assimilating human incarnate consciousness into Wetiko. I have watched during the 'Covid' hoax how many are becoming ever more profoundly attached to Wetiko's perceptual calling cards of aggressive response to any other point of view ('There is no other god but me'), psychopathic lack of compassion and empathy, and servile submission to the narrative and will of authority. Wetiko is the psychopaths *and* subservience to psychopaths. The Cult of Wetiko is

so anti-human because it is *not* human. It embarked on a mission to destroy human by targeting everything that it means to be human and to survive as human. 'Covid' is not the end, just a means to an end. The Cult with its Wetiko consciousness is seeking to change Earth systems, including the atmosphere, to suit them, not humans. The gathering bombardment of 5G alone from ground and space is dramatically changing The Field with which the five senses interact. There is so much more to come if we sit on our hands and hope it will all go away. It is not meant to go away. It is meant to get ever more extreme and we need to face that while we still can – just.

Carbon dioxide is the gas of life. Without that human is over. Kaput, gone, history. No natural world, no human. The Cult has created a cock and bull story about carbon dioxide and climate change to justify its reduction to the point where Gates and the ignoramus Biden 'climate chief' John Kerry want to suck it out of the atmosphere. Kerry wants to do this because his master Gates does. Wetikos have made the gas of life a demon with the usual support from the Wokers of Extinction Rebellion and similar organisations and the bewildered puppet-child that is Greta Thunberg who was put on the world stage by Klaus Schwab and the World Economic Forum. The name Extinction Rebellion is both ironic and as always Wetiko inversion. The gas that we need to survive must be reduced to save us from extinction. The most basic need of human is oxygen and we now have billions walking around in face nappies depriving body and brain of this essential requirement of human existence. More than that 5G at 60 gigahertz interacts with the oxygen molecule to reduce the amount of oxygen the body can absorb into the bloodstream. The obvious knock-on consequences of that for respiratory and cognitive problems and life itself need no further explanation. Psychopaths like Musk are assembling a global system of satellites to deluge the human atmosphere with this insanity. The man should be in jail. Here we have two most basic of human needs, oxygen and carbon dioxide, being dismantled.

Two others, water and food, are getting similar treatment with the United Nations Agendas 21 and 2030 – the Great Reset – planning to

centrally control all water and food supplies. People will not even own rain water that falls on their land. Food is affected at the most basic level by reducing carbon dioxide. We have genetic modification or GMO infiltrating the food chain on a mass scale, pesticides and herbicides polluting the air and destroying the soil. Freshwater fish that provide livelihoods for 60 million people and feed hundreds of millions worldwide are being 'pushed to the brink' according the conservationists while climate change is the only focus. Now we have Gates and Schwab wanting to dispense with current food sources all together and replace them with a synthetic version which the Wetiko Cult would control in terms of production and who eats and who doesn't. We have been on the Totalitarian Tiptoe to this for more than 60 years as food has become ever more processed and full of chemical shite to the point today when it's not natural food at all. As Dr Tom Cowan says: 'If it has a label don't eat it.' Bill Gates is now the biggest owner of farmland in the United States and he does nothing without an ulterior motive involving the Cult. Klaus Schwab wrote: 'To feed the world in the next 50 years we will need to produce as much food as was produced in the last 10,000 years … food security will only be achieved, however, if regulations on genetically modified foods are adapted to reflect the reality that gene editing offers a precise, efficient and safe method of improving crops.' Liar. People and the world are being targeted with aluminium through vaccines, chemtrails, food, drink cans, and endless other sources when aluminium has been linked to many health issues including dementia which is increasing year after year. Insects, bees and wildlife essential to the food chain are being deleted by pesticides, herbicides and radiation which 5G is dramatically increasing with 6G and 7G to come. The pollinating bee population is being devastated while wildlife including birds, dolphins and whales are having their natural radar blocked by the effects of ever-increasing radiation. In the summer windscreens used to be splattered with insects so numerous were they. It doesn't happen now. Where have they gone?

Synthetic everything

The Cult is introducing genetically-modified versions of trees, plants and insects including a Gates-funded project to unleash hundreds of millions of genetically-modified, lab-altered and patented male mosquitoes to mate with wild mosquitoes and induce genetic flaws that cause them to die out. Clinically-insane Gates-funded Japanese researchers have developed mosquitos that spread vaccine and are dubbed 'flying vaccinators'. Gates is funding the modification of weather patterns in part to sell the myth that this is caused by carbon dioxide and he's funding geoengineering of the skies to change the atmosphere. Some of this came to light with the Gates-backed plan to release tonnes of chalk into the atmosphere to 'deflect the Sun and cool the planet'. Funny how they do this while the heating effect of the Sun is not factored into climate projections focussed on carbon dioxide. The reason is that they want to reduce carbon dioxide (so don't mention the Sun), but at the same time they do want to reduce the impact of the Sun which is so essential to human life and health. I have mentioned the sun-cholesterol-vitamin D connection as they demonise the Sun with warnings about skin cancer (caused by the chemicals in sun cream they tell you to splash on). They come from the other end of the process with statin drugs to reduce cholesterol that turns sunlight into vitamin D. A lack of vitamin D leads to a long list of health effects and how vitamin D levels must have fallen with people confined to their homes over 'Covid'. Gates is funding other forms of geoengineering and most importantly chemtrails which are dropping heavy metals, aluminium and self-replicating nanotechnology onto the Earth which is killing the natural world. See *Everything You Need To Know, But Have Never Been Told* for the detailed background to this.

Every human system is being targeted for deletion by a force that's not human. The Wetiko Cult has embarked on the process of transforming the human body from biological to synthetic biological as I have explained. Biological is being replaced by the artificial and synthetic – Archontic 'countermimicry' – right across human society. The plan eventually is to dispense with the human body altogether

and absorb human consciousness – which it wouldn't really be by then – into cyberspace (the simulation which is Wetiko/Yaldabaoth). Preparations for that are already happening if people would care to look. The alternative media rightly warns about globalism and 'the globalists', but this is far bigger than that and represents the end of the human race as we know it. The 'bad copy' of prime reality that Gnostics describe was a bad copy of harmony, wonder and beauty to start with before Wetiko/Yaldabaoth set out to change the simulated 'copy' into something very different. The process was slow to start with. Entrapped humans in the simulation timeline were not technologically aware and they had to be brought up to intellectual speed while being suppressed spiritually to the point where they could build their own prison while having no idea they were doing so. We have now reached that stage where technological intellect has the potential to destroy us and that's why events are moving so fast. Central American shaman Don Juan Matus said:

> Think for a moment, and tell me how you would explain the contradictions between the intelligence of man the engineer and the stupidity of his systems of belief, or the stupidity of his contradictory behaviour. Sorcerers believe that the predators have given us our systems of beliefs, our ideas of good and evil; our social mores. They are the ones who set up our dreams of success or failure. They have given us covetousness, greed, and cowardice. It is the predator who makes us complacent, routinary, and egomaniacal.
>
> In order to keep us obedient and meek and weak, the predators engaged themselves in a stupendous manoeuvre – stupendous, of course, from the point of view of a fighting strategist; a horrendous manoeuvre from the point of those who suffer it. They gave us their mind. The predators' mind is baroque, contradictory, morose, filled with the fear of being discovered any minute now.

For 'predators' see Wetiko, Archons, Yaldabaoth, Jinn, and all the other versions of the same phenomenon in cultures and religions all over the world. The theme is always the same because it's true and it's real. We have reached the point where we have to deal with it. The question is – how?

Don't fight – walk away

I thought I'd use a controversial subheading to get things moving in terms of our response to global fascism. What do you mean 'don't fight'? What do you mean 'walk away'? We've got to fight. We can't walk away. Well, it depends what we mean by fight and walk away. If fighting means physical combat we are playing Wetiko's game and falling for its trap. It wants us to get angry, aggressive, and direct hate and hostility at the enemy we think we must fight. Every war, every battle, every conflict, has been fought with Wetiko leading both sides. It's what it does. Wetiko wants a fight, anywhere, any place. Just hit me, son, so I can hit you back. Wetiko hits Wetiko and Wetiko hits Wetiko in return. I am very forthright as you can see in exposing Wetikos of the Cult, but I don't hate them. I refuse to hate them. It's what they want. What you hate you become. What you *fight* you become. Wokers, 'anti-haters' and 'anti-fascists' prove this every time they reach for their keyboards or don their balaclavas. By walk away I mean to disengage from Wetiko which includes ceasing to cooperate with its tyranny. Paul Levy says of Wetiko:

> The way to 'defeat' evil is not to try to destroy it (for then, in playing evil's game, we have already lost), but rather, to find the invulnerable place within ourselves where evil is unable to vanquish us – this is to truly 'win' our battle with evil.

Wetiko is everywhere in human society and it's been on steroids since the 'Covid' hoax. Every shouting match over wearing masks has Wetiko wearing a mask and Wetiko not wearing one. It's an electrical circuit of push and resist, push and resist, with Wetiko pushing *and* resisting. Each polarity is Wetiko empowering itself. Dictionary definitions of 'resist' include 'opposing, refusing to accept or comply with' and the word to focus on is 'opposing'. What form does this take – setting police cars alight or 'refusing to accept or comply with'? The former is Wetiko opposing Wetiko while the other points the way forward. This is the difference between those aggressively demanding that government fascism must be obeyed who stand in stark contrast to the great majority of Pushbackers. We saw this clearly with a march by thousands of Pushbackers against lockdown in London followed days later by a Woker-hijacked

protest in Bristol in which police cars were set on fire. Masks were virtually absent in London and widespread in Bristol. Wetiko wants lockdown on every level of society and infuses its aggression to police it through its unknowing stooges. Lockdown protesters are the ones with the smiling faces and the hugs, The two blatantly obvious states of being – getting more obvious by the day – are the result of Wokers and their like becoming ever more influenced by the simulation Field of Wetiko and Pushbackers ever more influenced by The Field of a far higher vibration beyond the simulation. Wetiko can't invade the heart which is where most lockdown opponents are coming from. It's the heart that allows them to see through the lies to the truth in ways I will be highlighting.

Renegade Minds know that calmness is the place from which wisdom comes. You won't find wisdom in a hissing fit and wisdom is what we need in abundance right now. Calmness is not weakness – you don't have to scream at the top of your voice to be strong. Calmness is indeed a sign of strength. 'No' means I'm not doing it. *NOOOO!!!* doesn't mean you're not doing it even more. Volume does not advance 'No – I'm not doing it'. You are just not doing it. Wetiko possessed and influenced don't know how to deal with that. Wetiko wants a fight and we should not give it one. What it needs more than anything is our *cooperation* and we should not give that either. Mass rallies and marches are great in that they are a visual representation of feeling, but if it ends there they are irrelevant. You demand that Wetikos act differently? Well, they're not going to are they? They are Wetikos. We don't need to waste our time demanding that something doesn't happen when that will make no difference. We need to delete the means that *allows* it to happen. This, invariably, is our cooperation. You can demand a child stop firing a peashooter at the dog or you can refuse to buy the peashooter. If you provide the means you are cooperating with the dog being smacked on the nose with a pea. How can the authorities enforce mask-wearing if millions in a country refuse? What if the 74 million Pushbackers that voted for Trump in 2020 refused to wear masks, close their businesses or stay in their homes. It would be unenforceable. The

few control the many through the compliance of the many and that's always been the dynamic be it 'Covid' regulations or the Roman Empire. I know people can find it intimidating to say no to authority or stand out in a crowd for being the only one with a face on display; but it has to be done or it's over. I hope I've made clear in this book that where this is going will be far more intimidating than standing up now and saying 'No' – I will not cooperate with my own enslavement and that of my children. There might be consequences for some initially, although not so if enough do the same. The question that must be addressed is what is going to happen if we don't? It is time to be strong and unyieldingly so. No means no. Not here and there, but *everywhere* and *always*. I have refused to wear a mask and obey all the other nonsense. I will not comply with tyranny. I repeat: Fascism is not imposed by fascists – there are never enough of them. Fascism is imposed by the population acquiescing to fascism. *I will not do it*. I will die first, or my body will. Living meekly under fascism is a form of death anyway, the death of the spirit that Martin Luther King described.

Making things happen

We must not despair. This is not over till it's over and it's far from that. The 'fat lady' must refuse to sing. The longer the 'Covid' hoax has dragged on and impacted on more lives we have seen an awakening of phenomenal numbers of people worldwide to the realisation that what they have believed all their lives is not how the world really is. Research published by the system-serving University of Bristol and King's College London in February, 2021, concluded: 'One in every 11 people in Britain say they trust David Icke's take on the coronavirus pandemic.' It will be more by now and we have gathering numbers to build on. We must urgently progress from seeing the scam to ceasing to cooperate with it. Prominent German lawyer Reiner Fuellmich, also licenced to practice law in America, is doing a magnificent job taking the legal route to bring the psychopaths to justice through a second Nuremberg tribunal for crimes against humanity. Fuellmich has an impressive record of

beating the elite in court and he formed the German Corona Investigative Committee to pursue civil charges against the main perpetrators with a view to triggering criminal charges. Most importantly he has grasped the foundation of the hoax – the PCR test not testing for the 'virus' – and Christian Drosten is therefore on his charge sheet along with Gates frontman Tedros at the World Health Organization. Major players must be not be allowed to inflict their horrors on the human race without being brought to book. A life sentence must follow for Bill Gates and the rest of them. A group of researchers has also indicted the government of Norway for crimes against humanity with copies sent to the police and the International Criminal Court. The lawsuit cites participation in an internationally-planned false pandemic and violation of international law and human rights, the European Commission's definition of human rights by coercive rules, Nuremberg and Hague rules on fundamental human rights, and the Norwegian constitution. We must take the initiative from hereon and not just complain, protest and react.

There are practical ways to support vital mass non-cooperation. Organising in numbers is one. Lockdown marches in London in the spring in 2021 were mass non-cooperation that the authorities could not stop. There were too many people. Hundreds of thousands walked the London streets in the centre of the road for mile after mile while the Face-Nappies could only look on. They were determined, but calm, and just *did it* with no histrionics and lots of smiles. The police were impotent. Others are organising group shopping without masks for mutual support and imagine if that was happening all over. Policing it would be impossible. If the store refuses to serve people in these circumstances they would be faced with a long line of trolleys full of goods standing on their own and everything would have to be returned to the shelves. How would they cope with that if it kept happening? I am talking here about moving on from complaining to being pro-active; from watching things happen to making things happen. I include in this our relationship with the police. The behaviour of many Face-Nappies

has been disgraceful and anyone who thinks they would never find concentration camp guards in the 'enlightened' modern era have had that myth busted big-time. The period and setting may change – Wetikos never do. I watched film footage from a London march in which a police thug viciously kicked a protestor on the floor who had done nothing. His fellow Face-Nappies stood in a ring protecting him. What he did was a criminal assault and with a crowd far outnumbering the police this can no longer be allowed to happen unchallenged. I get it when people chant 'shame on you' in these circumstances, but that is no longer enough. They *have* no shame those who do this. Crowds needs to start making a citizen's arrest of the police who commit criminal offences and brutally attack innocent people and defenceless women. A citizen's arrest can be made under section 24A of the UK Police and Criminal Evidence (PACE) Act of 1984 and you will find something similar in other countries. I prefer to call it a Common Law arrest rather than citizen's for reasons I will come to shortly. Anyone can arrest a person committing an indictable offence or if they have reasonable grounds to suspect they are committing an indictable offence. On both counts the attack by the police thug would have fallen into this category. A citizen's arrest can be made to stop someone:

- Causing physical injury to himself or any other person
- Suffering physical injury
- Causing loss of or damage to property
- Making off before a constable can assume responsibility for him

A citizen's arrest may also be made to prevent a breach of the peace under Common Law and if they believe a breach of the peace will happen or anything related to harm likely to be done or already done in their presence. This is the way to go I think – the Common Law version. If police know that the crowd and members of the public will no longer be standing and watching while they commit

their thuggery and crimes they will think twice about acting like Brownshirts and Blackshirts.

Common Law – common sense

Mention of Common Law is very important. Most people think the law is the law as in one law. This is not the case. There are two bodies of law, Common Law and Statute Law, and they are not the same. Common Law is founded on the simple premise of do no harm. It does not recognise victimless crimes in which no harm is done while Statute Law does. There is a Statute Law against almost everything. So what is Statute Law? Amazingly it's the law of the *sea* that was brought ashore by the Cult to override the law of the land which is Common Law. They had no right to do this and as always they did it anyway. They had to. They could not impose their will on the people through Common Law which only applies to do no harm. How could you stitch up the fine detail of people's lives with that? Instead they took the law of the sea, or Admiralty Law, and applied it to the population. Statute Law refers to all the laws spewing out of governments and their agencies including all the fascist laws and regulations relating to 'Covid'. The key point to make is that Statute Law is *contract law*. It only applies between *contracting* corporations. Most police officers don't even know this. They have to be kept in the dark, too. Long ago when merchants and their sailing ships began to trade with different countries a contractual law was developed called Admiralty Law and other names. Again it only applied to *contracts* agreed between *corporate* entities. If there is no agreed contract the law of the sea had no jurisdiction *and that still applies to its new alias of Statute Law*. The problem for the Cult when the law of the sea was brought ashore was an obvious one. People were not corporations and neither were government entities. To overcome the latter they made governments and all associated organisations corporations. All the institutions are *private corporations* and I mean governments and their agencies, local councils, police, courts, military, US states, the whole lot. Go to the

Dun and Bradstreet corporate listings website for confirmation that they are all corporations. You are arrested by a private corporation called the police by someone who is really a private security guard and they take you to court which is another private corporation. Neither have jurisdiction over you unless you consent and *contract* with them. This is why you hear the mantra about law enforcement policing by *consent* of the people. In truth the people 'consent' only in theory through monumental trickery.

Okay, the Cult overcame the corporate law problem by making governments and institutions corporate entities; but what about people? They are not corporations are they? Ah ... well in a sense, and *only* a sense, they are. Not people exactly – the illusion of people. The Cult creates a corporation in the name of everyone at the time that their birth certificate is issued. Note birth/ *berth* certificate and when you go to court under the law of the sea on land you stand in a *dock*. These are throwbacks to the origin. My Common Law name is David Vaughan Icke. The name of the corporation created by the government when I was born is called Mr David Vaughan Icke usually written in capitals as MR DAVID VAUGHAN ICKE. That is not me, the living, breathing man. It is a fictitious corporate entity. The trick is to make you think that David Vaughan Icke and MR DAVID VAUGHAN ICKE are the same thing. *They are not*. When police charge you and take you to court they are prosecuting the corporate entity and not the living, breathing, man or woman. They have to trick you into identifying as the corporate entity and contracting with them. Otherwise they have no jurisdiction. They do this through a language known as legalese. Lawful and legal are not the same either. Lawful relates to Common Law and legal relates to Statute Law. Legalese is the language of Statue Law which uses terms that mean one thing to the public and another in legalese. Notice that when a police officer tells someone why they are being charged he or she will say at the end: 'Do you understand?' To the public that means 'Do you comprehend?' In legalese it means 'Do you stand under me?' Do you stand under my authority? If you say

yes to the question you are unknowingly agreeing to give them jurisdiction over you in a contract between two corporate entities.

This is a confidence trick in every way. Contracts have to be agreed between informed parties and if you don't know that David Vaughan Icke is agreeing to be the corporation MR DAVID VAUGHAN ICKE you cannot knowingly agree to contract. They are deceiving you and another way they do this is to ask for proof of identity. You usually show them a driving licence or other document on which your corporate name is written. In doing so you are accepting that you are that corporate entity when you are not. Referring to yourself as a 'person' or 'citizen' is also identifying with your corporate fiction which is why I made the Common Law point about the citizen's arrest. If you are approached by a police officer you identify yourself immediately as a living, breathing, man or woman and say 'I do not consent, I do not contract with you and I do not understand' or stand under their authority. I have a Common Law birth certificate as a living man and these are available at no charge from commonlawcourt.com. Businesses registered under the Statute Law system means that its laws apply. There are, however, ways to run a business under Common Law. Remember all 'Covid' laws and regulations are Statute Law – the law of *contracts* and you do not have to contract. This doesn't mean that you can kill someone and get away with it. Common Law says do no harm and that applies to physical harm, financial harm etc. Police are employees of private corporations and there needs to be a new system of non-corporate Common Law constables operating outside the Statute Law system. If you go to davidicke.com and put Common Law into the search engine you will find videos that explain Common Law in much greater detail. It is definitely a road we should walk.

With all my heart

I have heard people say that we are in a spiritual war. I don't like the term 'war' with its Wetiko dynamic, but I know what they mean. Sweep aside all the bodily forms and we are in a situation in which two states of consciousness are seeking very different realities.

Wetiko wants upheaval, chaos, fear, suffering, conflict and control. The other wants love, peace, harmony, fairness and freedom. That's where we are. We should not fall for the idea that Wetiko is all-powerful and there's nothing we can do. Wetiko is not all-powerful. It's a joke, pathetic. It doesn't have to be, but it has made that choice for now. A handful of times over the years when I have felt the presence of its frequency I have allowed it to attach briefly so I could consciously observe its nature. The experience is not pleasant, the energy is heavy and dark, but the ease with which you can kick it back out the door shows that its real power is in persuading us that it has power. It's all a con. Wetiko is a con. It's a trickster and not a power that can control us if we unleash our own. The con is founded on manipulating humanity to give its power to Wetiko which recycles it back to present the illusion that it has power when its power is *ours* that we gave away. This happens on an energetic level and plays out in the world of the seen as humanity giving its power to Wetiko authority which uses that power to control the population when the power is only the power the population has handed over. How could it be any other way for billions to be controlled by a relative few? I have had experiences with people possessed by Wetiko and again you can kick its arse if you do it with an open heart. Oh yes – the *heart* which can transform the world of perceived 'matter'.

We are receiver-transmitters and processors of information, but what information and where from? Information is processed into perception in three main areas – the brain, the heart and the belly. These relate to thinking, knowing, and emotion. Wetiko wants us to be head and belly people which means we think within the confines of the Matrix simulation and low-vibrational emotional reaction scrambles balance and perception. A few minutes on social media and you see how emotion is the dominant force. Woke is all emotion and is therefore thought-free and fact-free. Our heart is something different. It *knows* while the head *thinks* and has to try to work it out because it doesn't know. The human energy field has seven prime vortexes which connect us with wider reality (Fig 23). Chakra means

'wheels of light' in the Sanskrit language of ancient India. The main ones are: The crown chakra on top of the head; brow (or 'third eye') chakra in the centre of the forehead; throat chakra; heart chakra in the centre of the chest; solar plexus chakra below the sternum; sacral chakra beneath the navel; and base chakra at the bottom of the spine. Each one has a particular function or functions. We feel anxiety and nervousness in the belly where the sacral chakra is located and this processes emotion that can affect the colon to give people 'the shits' or make them 'shit scared' when they are nervous. Chakras all play an important role, but the Mr and Mrs Big is the heart chakra which sits at the centre of the seven, above the chakras that connect us to the 'physical' and below those that connect with higher realms (or at least should). Here in the heart chakra we feel love, empathy and compassion – 'My heart goes out to you'. Those with closed hearts become literally 'heart-less' in their attitudes and behaviour (see Bill Gates). Native Americans portrayed Wetiko with what Paul Levy calls a 'frigid, icy heart, devoid of mercy' (see Bill Gates).

Figure 23: The chakra system which interpenetrates the human energy field. The heart chakra is the governor – or should be.

Wetiko trembles at the thought of heart energy which it cannot infiltrate. The frequency is too high. What it seeks to do instead is close the heart chakra vortex to block its perceptual and energetic influence. Psychopaths have 'hearts of stone' and emotionally-damaged people have 'heartache' and 'broken hearts'. The astonishing amount of heart disease is related to heart chakra

disruption with its fundamental connection to the 'physical' heart. Dr Tom Cowan has written an outstanding book challenging the belief that the heart is a pump and making the connection between the 'physical' and spiritual heart. Rudolph Steiner who was way ahead of his time said the same about the fallacy that the heart is a pump. *What*? The heart is not a pump? That's crazy, right? Everybody knows that. Read Cowan's *Human Heart, Cosmic Heart* and you will realise that the very idea of the heart as a pump is ridiculous when you see the evidence. How does blood in the feet so far from the heart get pumped horizontally up the body by the heart?? Cowan explains in the book the real reason why blood moves as it does. Our 'physical' heart is used to symbolise love when the source is really the heart vortex or spiritual heart which is our most powerful energetic connection to 'out there' expanded consciousness. That's why we feel *knowing* – intuitive knowing – in the centre of the chest. Knowing doesn't come from a process of thoughts leading to a conclusion. It is there in an instant all in one go. Our heart knows because of its connection to levels of awareness that *do* know. This is the meaning and source of intuition – intuitive *knowing*.

For the last more than 30 years of uncovering the global game and the nature of reality my heart has been my constant antenna for truth and accuracy. An American intelligence insider once said that I had quoted a disinformer in one of my books and yet I had only quoted the part that was true. He asked: 'How do you do that?' By using my heart antenna was the answer and anyone can do it. Heart-centred is how we are meant to be. With a closed heart chakra we withdraw into a closed mind and the bubble of five-sense reality. If you take a moment to focus your attention on the centre of your chest, picture a spinning wheel of light and see it opening and expanding. You will feel it happening, too, and perceptions of the heart like joy and love as the heart impacts on the mind as they interact. The more the chakra opens the more you will feel expressions of heart consciousness and as the process continues, and becomes part of you, insights and knowings will follow. An open

heart is connected to that level of awareness that knows all is *One*. You will see from its perspective that the fault-lines that divide us are only illusions to control us. An open heart does not process the illusions of race, creed and sexuality except as brief experiences for a consciousness that is all. Our heart does not see division, only unity (Figs 24 and 25). There's something else, too. Our hearts love to laugh. Mark Twain's quote that says 'The human race has one really effective weapon, and that is laughter' is really a reference to the heart which loves to laugh with the joy of knowing the true nature of infinite reality and that all the madness of human society is an illusion of the mind. Twain also said: 'Against the assault of laughter nothing can stand.' This is so true of Wetiko and the Cult. Their insecurity demands that they be taken seriously and their power and authority acknowledged and feared. We should do nothing of the sort. We should not get aggressive or fearful which their insecurity so desires. We should laugh in their face. Even in their no-face as police come over in their face-nappies and expect to be taken seriously. They don't take themselves seriously looking like that so why should we? Laugh in the face of intimidation. Laugh in the face of tyranny. You will see by its reaction that you have pressed all of its buttons. Wetiko does not know what to do in the face of laughter or when its targets refuse to concede their joy to fear. We have seen many examples during the 'Covid' hoax when people have expressed their energetic power and the string puppets of Wetiko retreat with their tail limp between their knees. Laugh – the world is bloody mad after all and if it's a choice between laughter and tears I know which way I'm going.

Figure 24: Head consciousness without the heart sees division and everything apart from everything else.

Figure 25: Heart consciousness sees everything as One.

'Vaccines' and the soul

The foundation of Wetiko/Archon control of humans is the separation of incarnate five-sense mind from the infinite 'I' and closing the heart chakra where the True 'I' lives during a human life. The goal has been to achieve complete separation in both cases. I was interested therefore to read an account by a French energetic healer of what she said she experienced with a patient who had been given the 'Covid' vaccine. Genuine energy healers can sense information and consciousness fields at different levels of being which are referred to as 'subtle bodies'. She described treating the patient who later returned after having, without the healer's knowledge, two doses of the 'Covid vaccine'. The healer said:

I noticed immediately the change, very heavy energy emanating from [the] subtle bodies. The scariest thing was when I was working on the heart chakra, I connected with her soul: it was detached from the physical body, it had no contact and it was, as if it was floating in a state of total confusion: a damage to the consciousness that loses contact with the physical body, i.e. with our biological machine, there is no longer any communication between them.

I continued the treatment by sending light to the heart chakra, the soul of the person, but it seemed that the soul could no longer receive any light, frequency or energy. It was a very powerful experience for me. Then I understood that this substance is indeed used to detach consciousness so that this consciousness can no longer interact through this body that it possesses in life, where there is no longer any contact, no frequency, no light, no more energetic balance or mind.

This would create a human that is rudderless and at the extreme almost zombie-like operating with a fractional state of consciousness at the mercy of Wetiko. I was especially intrigued by what the healer said in the light of the prediction by the highly-informed Rudolf Steiner more than a hundred years ago. He said:

In the future, we will eliminate the soul with medicine. Under the pretext of a 'healthy point of view', there will be a vaccine by which the human body will be treated as soon as possible directly at birth, so that the human being cannot develop the thought of the existence of soul and Spirit. To materialistic doctors will be entrusted the task of removing the soul of humanity.

As today, people are vaccinated against this disease or that disease, so in the future, children will be vaccinated with a substance that can be produced precisely in such a way that people, thanks to this vaccination, will be immune to being subjected to the 'madness' of spiritual life. He would be extremely smart, but he would not develop a conscience, and that is the true goal of some materialistic circles.

Steiner said the vaccine would detach the physical body from the etheric body (subtle bodies) and 'once the etheric body is detached the relationship between the universe and the etheric body would become extremely unstable, and man would become an automaton'. He said 'the physical body of man must be polished on this Earth by spiritual will – so the vaccine becomes a kind of arymanique (Wetiko) force' and 'man can no longer get rid of a given materialistic feeling'. Humans would then, he said, become 'materialistic of constitution and can no longer rise to the spiritual'. I have been writing for years about DNA being a receiver-transmitter of information that connects us to other levels of reality and these 'vaccines' changing DNA can be likened to changing an antenna and what it can transmit and receive. Such a disconnection would clearly lead to changes in personality and perception. Steiner further predicted the arrival of AI. Big Pharma 'Covid vaccine' makers, expressions of Wetiko, are testing their DNA-manipulating evil on children as I write with a view to giving the 'vaccine' to babies. If it's a soul-body disconnector – and I say that it is or can be – every child would be disconnected from 'soul' at birth and the 'vaccine' would create a closed system in which spiritual guidance from the greater self would play no part. This has been the ambition of Wetiko all

along. A Pentagon video from 2005 was leaked of a presentation explaining the development of vaccines to change behaviour by their effect on the brain. Those that believe this is not happening with the 'Covid' genetically-modifying procedure masquerading as a 'vaccine' should make an urgent appointment with Naivety Anonymous. Klaus Schwab wrote in 2018:

> Neurotechnologies enable us to better influence consciousness and thought and to understand many activities of the brain. They include decoding what we are thinking in fine levels of detail through new chemicals and interventions that can influence our brains to correct for errors or enhance functionality.

The plan is clear and only the heart can stop it. With every heart that opens, every mind that awakens, Wetiko is weakened. Heart and love are far more powerful than head and hate and so nothing like a majority is needed to turn this around.

Beyond the Phantom

Our heart is the prime target of Wetiko and so it must be the answer to Wetiko. We *are* our heart which is part of one heart, the infinite heart. Our heart is where the true self lives in a human life behind firewalls of five-sense illusion when an imposter takes its place – *Phantom Self*; but our heart waits patiently to be set free any time we choose to see beyond the Phantom, beyond Wetiko. A Wetikoed Phantom Self can wreak mass death and destruction while the love of forever is locked away in its heart. The time is here to unleash its power and let it sweep away the fear and despair that is Wetiko. Heart consciousness does not seek manipulated, censored, advantage for its belief or religion, its activism and desires. As an expression of the One it treats all as One with the same rights to freedom and opinion. Our heart demands fairness for itself no more than for others. From this unity of heart we can come together in mutual support and transform this Wetikoed world into what reality is meant to be – a place of love, joy, happiness, fairness, justice and freedom. Wetiko has another agenda and that's why the world is as

it is, but enough of this nonsense. Wetiko can't stay where hearts are open and it works so hard to keep them closed. Fear is its currency and its food source and love in its true sense has no fear. Why would love have fear when it knows it is *All That Is, Has Been, And Ever Can Be* on an eternal exploration of all possibility? Love in this true sense is not the physical attraction that passes for love. This can be an expression of it, yes, but Infinite Love, a love without condition, goes far deeper to the core of all being. It *is* the core of all being. Infinite realty was born from love beyond the illusions of the simulation. Love infinitely expressed is the knowing that all is One and the swiftly-passing experience of separation is a temporary hallucination. You cannot disconnect from Oneness; you can only *perceive* that you have and withdraw from its influence. This is the most important of all perception trickery by the mind parasite that is Wetiko and the foundation of all its potential for manipulation.

If we open our hearts, open the sluice gates of the mind, and redefine self-identity amazing things start to happen. Consciousness expands or contracts in accordance with self-identity. When true self is recognised as infinite awareness and label self – Phantom Self – is seen as only a series of brief experiences life is transformed. Consciousness expands to the extent that self-identity expands and everything changes. You see unity, not division, the picture, not the pixels. From this we can play the long game. No more is an experience something in and of itself, but a fleeting moment in the eternity of forever. Suddenly people in uniform and dark suits are no longer intimidating. Doing what your heart knows to be right is no longer intimidating and consequences for those actions take on the same nature of a brief experience that passes in the blink of an infinite eye. Intimidation is all in the mind. Beyond the mind there is no intimidation.

An open heart does not consider consequences for what it knows to be right. To do so would be to consider not doing what it knows to be right and for a heart in its power that is never an option. The Renegade Mind is really the Renegade Heart. Consideration of consequences will always provide a getaway car for the mind and

the heart doesn't want one. What is right in the light of what we face today is to stop cooperating with Wetiko in all its forms and to do it without fear or compromise. You cannot compromise with tyranny when tyranny always demands more until it has everything. Life is your perception and you are your destiny. Change your perception and you change your life. Change collective perception and we change the world.

Come on people … One human family, One heart, One goal … *FREEEEEEDOM*!

We must settle for nothing less.

Postscript

The big scare story as the book goes to press is the 'Indian' variant and the world is being deluged with propaganda about the 'Covid catastrophe' in India which mirrors in its lies and misrepresentations what happened in Italy before the first lockdown in 2020.

The *New York Post* published a picture of someone who had 'collapsed in the street from Covid' in India in April, 2021, which was actually taken during a gas leak in May, 2020. Same old, same old. Media articles in mid-February were asking why India had been so untouched by 'Covid' and then as their vaccine rollout gathered pace the alleged 'cases' began to rapidly increase. Indian 'Covid vaccine' maker Bharat Biotech was funded into existence by the Bill and Melinda Gates Foundation (the pair announced their divorce in May, 2021, which is a pity because they so deserve each other). The Indian 'Covid crisis' was ramped up by the media to terrify the world and prepare people for submission to still more restrictions. The scam that worked the first time was being repeated only with far more people seeing through the deceit. Davidicke.com and Ickonic.com have sought to tell the true story of what is happening by talking to people living through the Indian nightmare which has nothing to do with 'Covid'. We posted a letter from 'Alisha' in Pune who told a very different story to government and media mendacity. She said scenes of dying people and overwhelmed hospitals were designed to hide what was really happening – genocide and starvation. Alisha said that millions had already died of starvation during the ongoing lockdowns while government and media were lying and making it look like the 'virus':

Restaurants, shops, gyms, theatres, basically everything is shut. The cities are ghost towns. Even so-called 'essential' businesses are only open till 11am in the morning. You basically have just an hour to buy food and then your time is up.

Inter-state travel and even inter-district travel is banned. The cops wait at all major crossroads to question why you are traveling outdoors or to fine you if you are not wearing a mask.

The medical community here is also complicit in genocide, lying about hospitals being full and turning away people with genuine illnesses, who need immediate care. They have even created a shortage of oxygen cylinders.

This is the classic Cult modus operandi played out in every country. Alisha said that people who would not have a PCR test not testing for the 'virus' were being denied hospital treatment. She said the people hit hardest were migrant workers and those in rural areas. Most businesses employed migrant workers and with everything closed there were no jobs, no income and no food. As a result millions were dying of starvation or malnutrition. All this was happening under Prime Minister Narendra Modi, a 100-percent asset of the Cult, and it emphasises yet again the scale of pure anti-human evil we are dealing with. Australia banned its people from returning home from India with penalties for trying to do so of up to five years in jail and a fine of £37,000. The manufactured 'Covid' crisis in India was being prepared to justify further fascism in the West. Obvious connections could be seen between the Indian 'vaccine' programme and increased 'cases' and this became a common theme. The Seychelles, the most per capita 'Covid vaccinated' population in the world, went back into lockdown after a 'surge of cases'.

Long ago the truly evil Monsanto agricultural biotechnology corporation with its big connections to Bill Gates devastated Indian farming with genetically-modified crops. Human rights activist Gurcharan Singh highlighted the efforts by the Indian government to complete the job by destroying the food supply to hundreds of millions with 'Covid' lockdowns. He said that 415 million people at the bottom of the disgusting caste system (still going whatever they say) were below the poverty line and struggled to feed themselves every year. Now the government was imposing lockdown at just the

time to destroy the harvest. This deliberate policy was leading to mass starvation. People may reel back at the suggestion that a government would do that, but Wetiko-controlled 'leaders' are capable of any level of evil. In fact what is described in India is in the process of being instigated worldwide. The food chain and food supply are being targeted at every level to cause world hunger and thus control. Bill Gates is not the biggest owner of farmland in America for no reason and destroying access to food aids both the depopulation agenda and the plan for synthetic 'food' already being funded into existence by Gates. Add to this the coming hyper-inflation from the suicidal creation of fake 'money' in response to 'Covid' and the breakdown of container shipping systems and you have a cocktail that can only lead one way and is meant to. The Cult plan is to crash the entire system to 'build back better' with the Great Reset.

'Vaccine' transmission

Reports from all over the world continue to emerge of women suffering menstrual and fertility problems after having the fake 'vaccine' and of the non-'vaccinated' having similar problems when interacting with the 'vaccinated'. There are far too many for 'coincidence' to be credible. We've had menopausal women getting periods, others having periods stop or not stopping for weeks, passing clots, sometimes the lining of the uterus, breast irregularities, and miscarriages (which increased by 400 percent in parts of the United States). Non-'vaccinated' men and children have suffered blood clots and nose bleeding after interaction with the 'vaccinated'. Babies have died from the effects of breast milk from a 'vaccinated' mother. Awake doctors – the small minority – speculated on the cause of non-'vaccinated' suffering the same effects as the 'vaccinated'. Was it nanotechnology in the synthetic substance transmitting frequencies or was it a straight chemical bioweapon that was being transmitted between people? I am not saying that some kind of chemical transmission is not one possible answer, but the foundation of all that the Cult does is frequency and

this is fertile ground for understanding how transmission can happen. American doctor Carrie Madej, an internal medicine physician and osteopath, has been practicing for the last 20 years, teaching medical students, and she says attending different meetings where the agenda for humanity was discussed. Madej, who operates out of Georgia, did not dismiss other possible forms of transmission, but she focused on frequency in search of an explanation for transmission. She said the Moderna and Pfizer 'vaccines' contained nano-lipid particles as a key component. This was a brand new technology never before used on humanity. 'They're using a nanotechnology which is pretty much little tiny computer bits … nanobots or hydrogel.' Inside the 'vaccines' was 'this sci-fi kind of substance' which suppressed immune checkpoints to get into the cell. I referred to this earlier as the 'Trojan horse' technique that tricks the cell into opening a gateway for the self-replicating synthetic material and while the immune system is artificially suppressed the body has no defences. Madej said the substance served many purposes including an on-demand ability to 'deliver the payload' and using the nano 'computer bits' as biosensors in the body. 'It actually has the ability to accumulate data from your body, like your breathing, your respiration, thoughts, emotions, all kinds of things.'

She said the technology obviously has the ability to operate through Wi-Fi and transmit and receive energy, messages, frequencies or impulses. 'Just imagine you're getting this new substance in you and it can react to things all around you, the 5G, your smart device, your phones.' We had something completely foreign in the human body that had never been launched large scale at a time when we were seeing 5G going into schools and hospitals (plus the Musk satellites) and she believed the 'vaccine' transmission had something to do with this: '… if these people have this inside of them … it can act like an antenna and actually transmit it outwardly as well.' The synthetic substance produced its own voltage and so it could have that kind of effect. This fits with my own contention that the nano receiver-transmitters are designed to connect people to the

Smart Grid and break the receiver-transmitter connection to expanded consciousness. That would explain the French energy healer's experience of the disconnection of body from 'soul' with those who have had the 'vaccine'. The nanobots, self-replicating inside the body, would also transmit the synthetic frequency which could be picked up through close interaction by those who have not been 'vaccinated'. Madej speculated that perhaps it was 5G and increased levels of other radiation that was causing the symptoms directly although interestingly she said that non-'vaccinated' patients had shown improvement when they were away from the 'vaccinated' person they had interacted with. It must be remembered that you can control frequency and energy with your mind and you can consciously create energetic barriers or bubbles with the mind to stop damaging frequencies from penetrating your field. American paediatrician Dr Larry Palevsky said the 'vaccine' was not a 'vaccine' and was never designed to protect from a 'viral' infection. He called it 'a massive, brilliant propaganda of genocide' because they didn't have to inject everyone to get the result they wanted. He said the content of the jabs was able to infuse any material into the brain, heart, lungs, kidneys, liver, sperm and female productive system. 'This is genocide; this is a weapon of mass destruction.' At the same time American colleges were banning students from attending if they didn't have this life-changing and potentially life-ending 'vaccine'. Class action lawsuits must follow when the consequences of this college fascism come to light. As the book was going to press came reports about fertility effects on sperm in 'vaccinated' men which would absolutely fit with what I have been saying and hospitals continued to fill with 'vaccine' reactions. Another question is what about transmission via blood transfusions? The NHS has extended blood donation restrictions from seven days after a 'Covid vaccination' to 28 days after even a sore arm reaction.

I said in the spring of 2020 that the then touted 'Covid vaccine' would be ongoing each year like the flu jab. A year later Pfizer CEO, the appalling Albert Bourla, said people would 'likely' need a 'booster dose' of the 'vaccine' within 12 months of getting 'fully

vaccinated' and then a yearly shot. 'Variants will play a key role', he said confirming the point. Johnson & Johnson CEO Alex Gorsky also took time out from his 'vaccine' disaster to say that people may need to be vaccinated against 'Covid-19' each year. UK Health Secretary, the psychopath Matt Hancock, said additional 'boosters' would be available in the autumn of 2021. This is the trap of the 'vaccine passport'. The public will have to accept every last 'vaccine' they introduce, including for the fake 'variants', or it would cease to be valid. The only other way in some cases would be continuous testing with a test not testing for the 'virus' and what is on the swabs constantly pushed up your noise towards the brain every time?

'Vaccines' changing behaviour

I mentioned in the body of the book how I believed we would see gathering behaviour changes in the 'vaccinated' and I am already hearing such comments from the non-'vaccinated' describing behaviour changes in friends, loved ones and work colleagues. This will only increase as the self-replicating synthetic material and nanoparticles expand in body and brain. An article in the *Guardian* in 2016 detailed research at the University of Virginia in Charlottesville which developed a new method for controlling brain circuits associated with complex animal behaviour. The method, dubbed 'magnetogenetics', involves genetically-engineering a protein called ferritin, which stores and releases iron, to create a magnetised substance – 'Magneto' – that can activate specific groups of nerve cells from a distance. This is claimed to be an advance on other methods of brain activity manipulation known as optogenetics and chemogenetics (the Cult has been developing methods of brain control for a long time). The ferritin technique is said to be non-invasive and able to activate neurons 'rapidly and reversibly'. In other words, human thought and perception. The article said that earlier studies revealed how nerve cell proteins 'activated by heat and mechanical pressure can be genetically engineered so that they become sensitive to radio waves and magnetic fields, by attaching them to an iron-storing protein called ferritin, or to inorganic

paramagnetic particles'. Sensitive to radio waves and magnetic fields? You mean like 5G, 6G and 7G? This is the human-AI Smart Grid hive mind we are talking about. The *Guardian* article said:

> … the researchers injected Magneto into the striatum of freely behaving mice, a deep brain structure containing dopamine-producing neurons that are involved in reward and motivation, and then placed the animals into an apparatus split into magnetised and non-magnetised sections.
>
> Mice expressing Magneto spent far more time in the magnetised areas than mice that did not, because activation of the protein caused the striatal neurons expressing it to release dopamine, so that the mice found being in those areas rewarding. This shows that Magneto can remotely control the firing of neurons deep within the brain, and also control complex behaviours.

Make no mistake this basic methodology will be part of the 'Covid vaccine' cocktail and using magnetics to change brain function through electromagnetic field frequency activation. The Pentagon is developing a 'Covid vaccine' using ferritin. Magnetics would explain changes in behaviour and why videos are appearing across the Internet as I write showing how magnets stick to the skin at the point of the 'vaccine' shot. Once people take these 'vaccines' anything becomes possible in terms of brain function and illness which will be blamed on 'Covid-19' and 'variants'. Magnetic field manipulation would further explain why the non-'vaccinated' are reporting the same symptoms as the 'vaccinated' they interact with and why those symptoms are reported to decrease when not in their company. Interestingly 'Magneto', a 'mutant', is a character in the Marvel Comic *X-Men* stories with the ability to manipulate magnetic fields and he believes that mutants should fight back against their human oppressors by any means necessary. The character was born Erik Lehnsherr to a Jewish family in Germany.

Cult-controlled courts

The European Court of Human Rights opened the door for mandatory 'Covid-19 vaccines' across the continent when it ruled in a Czech Republic dispute over childhood immunisation that legally

enforced vaccination could be 'necessary in a democratic society'. The 17 judges decided that compulsory vaccinations did not breach human rights law. On the face of it the judgement was so inverted you gasp for air. If not having a vaccine infused into your body is not a human right then what is? Ah, but they said human rights law which has been specifically written to delete all human rights at the behest of the state (the Cult). Article 8 of the European Convention on Human Rights relates to the right to a private life. The crucial word here is *'except'*:

> There shall be no interference by a public authority with the exercise of this right EXCEPT such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others [My emphasis].

No interference *except* in accordance with the law means there *are* no 'human rights' *except* what EU governments decide you can have at their behest. 'As is necessary in a democratic society' explains that reference in the judgement and 'in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others' gives the EU a coach and horses to ride through 'human rights' and scatter them in all directions. The judiciary is not a check and balance on government extremism; it is a vehicle to enforce it. This judgement was almost laughably predictable when the last thing the Cult wanted was a decision that went against mandatory vaccination. Judges rule over and over again to benefit the system of which they are a part. Vaccination disputes that come before them are invariably delivered in favour of doctors and authorities representing the view of the state which owns the judiciary. Oh, yes, and we have even had calls to stop putting 'Covid-19' on death certificates within 28 days of a 'positive test' because it is claimed the practice makes the 'vaccine' appear not to work. They are laughing at you.

The scale of madness, inhumanity and things to come was highlighted when those not 'vaccinated' for 'Covid' were refused evacuation from the Caribbean island of St Vincent during massive volcanic eruptions. Cruise ships taking residents to the safety of another island allowed only the 'vaccinated' to board and the rest were left to their fate. Even in life and death situations like this we see 'Covid' stripping people of their most basic human instincts and the insanity is even more extreme when you think that fake 'vaccine'-makers are not even claiming their body-manipulating concoctions stop 'infection' and 'transmission' of a 'virus' that doesn't exist. St Vincent Prime Minister Ralph Gonsalves said: 'The chief medical officer will be identifying the persons already vaccinated so that we can get them on the ship.' Note again the power of the chief medical officer who, like Whitty in the UK, will be answering to the World Health Organization. This is the Cult network structure that has overridden politicians who 'follow the science' which means doing what WHO-controlled 'medical officers' and 'science advisers' tell them. Gonsalves even said that residents who were 'vaccinated' after the order so they could board the ships would still be refused entry due to possible side effects such as 'wooziness in the head'. The good news is that if they were woozy enough in the head they could qualify to be prime minister of St Vincent.

Microchipping freedom

The European judgement will be used at some point to justify moves to enforce the 'Covid' DNA-manipulating procedure. Sandra Ro, CEO of the Global Blockchain Business Council, told a World Economic Forum event that she hoped 'vaccine passports' would help to 'drive forced consent and standardisation' of global digital identity schemes: 'I'm hoping with the desire and global demand for some sort of vaccine passport – so that people can get travelling and working again – [it] will drive forced consent, standardisation, and frankly, cooperation across the world.' The lady is either not very bright, or thoroughly mendacious, to use the term 'forced consent'.

You do not 'consent' if you are forced – you *submit*. She was describing what the plan has been all along and that's to enforce a digital identity on every human without which they could not function. 'Vaccine passports' are opening the door and are far from the end goal. A digital identity would allow you to be tracked in everything you do in cyberspace and this is the same technique used by Cult-owned China to enforce its social credit system of total control. The ultimate 'passport' is planned to be a microchip as my books have warned for nearly 30 years. Those nice people at the Pentagon working for the Cult-controlled Defense Advanced Research Projects Agency (DARPA) claimed in April, 2021, they have developed a microchip inserted under the skin to detect 'asymptomatic Covid-19 infection' before it becomes an outbreak and a 'revolutionary filter' that can remove the 'virus' from the blood when attached to a dialysis machine. The only problems with this are that the 'virus' does not exist and people transmitting the 'virus' with no symptoms is brain-numbing bullshit. This is, of course, not a ruse to get people to be microchipped for very different reasons. DARPA also said it was producing a one-stop 'vaccine' for the 'virus' and all 'variants'. One of the most sinister organisations on Planet Earth is doing this? Better have it then. These people are insane because Wetiko that possesses them is insane.

Researchers from the Salk Institute in California announced they have created an embryo that is part human and part monkey. My books going back to the 1990s have exposed experiments in top secret underground facilities in the United States where humans are being crossed with animal and non-human 'extraterrestrial' species. They are now easing that long-developed capability into the public arena and there is much more to come given we are dealing with psychiatric basket cases. Talking of which – Elon Musk's scientists at Neuralink trained a monkey to play Pong and other puzzles on a computer screen using a joystick and when the monkey made the correct move a metal tube squirted banana smoothie into his mouth which is the basic technique for training humans into unquestioning compliance. Two Neuralink chips were in the monkey's skull and

more than 2,000 wires 'fanned out' into its brain. Eventually the monkey played a video game purely with its brain waves. Psychopathic narcissist Musk said the 'breakthrough' was a step towards putting Neuralink chips into human skulls and merging minds with artificial intelligence. *Exactly*. This man is so dark and Cult to his DNA.

World Economic Fascism (WEF)

The World Economic Forum is telling you the plan by the statements made at its many and various events. Cult-owned fascist YouTube CEO Susan Wojcicki spoke at the 2021 WEF Global Technology Governance Summit (see the name) in which 40 governments and 150 companies met to ensure 'the responsible design and deployment of emerging technologies'. Orwellian translation: 'Ensuring the design and deployment of long-planned technologies will advance the Cult agenda for control and censorship.' Freedom-destroyer and Nuremberg-bound Wojcicki expressed support for tech platforms like hers to censor content that is 'technically legal but could be harmful'. Who decides what is 'harmful'? She does and they do. 'Harmful' will be whatever the Cult doesn't want people to see and we have legislation proposed by the UK government that would censor content on the basis of 'harm' no matter if the information is fair, legal and provably true. Make that *especially* if it is fair, legal and provably true. Wojcicki called for a global coalition to be formed to enforce content moderation standards through automated censorship. This is a woman and mega-censor so self-deluded that she shamelessly accepted a 'free expression' award – *Wojcicki* – in an event sponsored by her own *YouTube*. They have no shame and no self-awareness.

You know that 'Covid' is a scam and Wojcicki a Cult operative when YouTube is censoring medical and scientific opinion purely on the grounds of whether it supports or opposes the Cult 'Covid' narrative. Florida governor Ron DeSantis compiled an expert panel with four professors of medicine from Harvard, Oxford, and Stanford Universities who spoke against forcing children and

vaccinated people to wear masks. They also said there was no proof that lockdowns reduced spread or death rates of 'Covid-19'. Cult-gofer Wojcicki and her YouTube deleted the panel video 'because it included content that contradicts the consensus of local and global health authorities regarding the efficacy of masks to prevent the spread of Covid-19'. This 'consensus' refers to what the Cult tells the World Health Organization to say and the WHO tells 'local health authorities' to do. Wojcicki knows this, of course. The panellists pointed out that censorship of scientific debate was responsible for deaths from many causes, but Wojcicki couldn't care less. She would not dare go against what she is told and as a disgrace to humanity she wouldn't want to anyway. The UK government is seeking to pass a fascist 'Online Safety Bill' to specifically target with massive fines and other means non-censored video and social media platforms to make them censor 'lawful but harmful' content like the Cult-owned Facebook, Twitter, Google and YouTube. What is 'lawful but harmful' would be decided by the fascist Blair-created Ofcom.

Another WEF obsession is a cyber-attack on the financial system and this is clearly what the Cult has planned to take down the bank accounts of everyone – except theirs. Those that think they have enough money for the Cult agenda not to matter to them have got a big lesson coming if they continue to ignore what is staring them in the face. The World Economic Forum, funded by Gates and fronted by Klaus Schwab, announced it would be running a 'simulation' with the Russian government and global banks of just such an attack called Cyber Polygon 2021. What they simulate – as with the 'Covid' Event 201 – they plan to instigate. The WEF is involved in a project with the Cult-owned Carnegie Endowment for International Peace called the WEF-Carnegie Cyber Policy Initiative which seeks to merge Wall Street banks, 'regulators' (I love it) and intelligence agencies to 'prevent' (arrange and allow) a cyber-attack that would bring down the global financial system as long planned by those that control the WEF and the Carnegie operation. The Carnegie Endowment for International Peace sent an instruction to First World

War US President Woodrow Wilson not to let the war end before society had been irreversibly transformed.

The Wuhan lab diversion

As I close, the Cult-controlled authorities and lapdog media are systematically pushing 'the virus was released from the Wuhan lab' narrative. There are two versions – it happened by accident and it happened on purpose. Both are nonsense. The perceived existence of the never-shown-to-exist 'virus' is vital to sell the impression that there is actually an infective agent to deal with and to allow the endless potential for terrifying the population with 'variants' of a 'virus' that does not exist. The authorities at the time of writing are going with the 'by accident' while the alternative media is promoting the 'on purpose'. Cable news host Tucker Carlson who has questioned aspects of lockdown and 'vaccine' compulsion has bought the Wuhan lab story. 'Everyone now agrees' he said. Well, I don't and many others don't and the question is *why* does the system and its media suddenly 'agree'? When the media moves as one unit with a narrative it is always a lie – witness the hour by hour mendacity of the 'Covid' era. Why would this Cult-owned combination which has unleashed lies like machine gun fire suddenly 'agree' to tell the truth??

Much of the alternative media is buying the lie because it fits the conspiracy narrative, but it's the *wrong* conspiracy. The real conspiracy is that *there is no virus* and that is what the Cult is desperate to hide. The idea that the 'virus' was released by accident is ludicrous when the whole 'Covid' hoax was clearly long-planned and waiting to be played out as it was so fast in accordance with the Rockefeller document and Event 201. So they prepared everything in detail over decades and then sat around strumming their fingers waiting for an 'accidental' release from a bio-lab? *What??* It's crazy. Then there's the 'on purpose' claim. You want to circulate a 'deadly virus' and hide the fact that you've done so and you release it down the street from the highest-level bio-lab in China? I repeat – *What??*

You would release it far from that lab to stop any association being made. But, no, we'll do it in a place where the connection was certain to be made. Why would you need to scam 'cases' and 'deaths' and pay hospitals to diagnose 'Covid-19' if you had a real 'virus'? What are sections of the alternative media doing believing this crap? Where were all the mass deaths in Wuhan from a 'deadly pathogen' when the recovery to normal life after the initial propaganda was dramatic in speed? Why isn't the 'deadly pathogen' now circulating all over China with bodies in the street? Once again we have the technique of tell them what they want to hear and they will likely believe it. The alternative media has its 'conspiracy' and with Carlson it fits with his 'China is the danger' narrative over years. China *is* a danger as a global Cult operations centre, but not for this reason. The Wuhan lab story also has the potential to instigate conflict with China when at some stage the plan is to trigger a Problem-Reaction-Solution confrontation with the West. Question everything – *everything* – and especially when the media agrees on a common party line.

Third wave ... fourth wave ... fifth wave ...

As the book went into production the world was being set up for more lockdowns and a 'third wave' supported by invented 'variants' that were increasing all the time and will continue to do so in public statements and computer programs, but not in reality. India became the new Italy in the 'Covid' propaganda campaign and we were told to be frightened of the new 'Indian strain'. Somehow I couldn't find it within myself to do so. A document produced for the UK government entitled 'Summary of further modelling of easing of restrictions – Roadmap Step 2' declared that a third wave was inevitable (of course when it's in the script) and it would be the fault of children and those who refuse the health-destroying fake 'Covid vaccine'. One of the computer models involved came from the Cult-owned *Imperial College* and the other from Warwick University which I wouldn't trust to tell me the date in a calendar factory. The document states that both models presumed extremely high uptake

of the 'Covid vaccines' and didn't allow for 'variants'. The document states: 'The resurgence is a result of some people (mostly children) being ineligible for vaccination; others choosing not to receive the vaccine; and others being vaccinated but not perfectly protected.' The mendacity takes the breath away. Okay, blame those with a brain who won't take the DNA-modifying shots and put more pressure on children to have it as 'trials' were underway involving children as young as six months with parents who give insanity a bad name. Massive pressure is being put on the young to have the fake 'vaccine' and child age consent limits have been systematically lowered around the world to stop parents intervening. Most extraordinary about the document was its claim that the 'third wave' would be driven by 'the resurgence in both hospitalisations and deaths … dominated by *those that have received two doses of the vaccine*, comprising around 60-70% of the wave respectively'. The predicted peak of the 'third wave' suggested 300 deaths per day with 250 of them *fully 'vaccinated' people*. How many more lies do acquiescers need to be told before they see the obvious? Those who took the jab to 'protect themselves' are projected to be those who mostly get sick and die? So what's in the 'vaccine'? The document went on:

> It is possible that a summer of low prevalence could be followed by substantial increases in incidence over the following autumn and winter. Low prevalence in late summer should not be taken as an indication that SARS-CoV-2 has retreated or that the population has high enough levels of immunity to prevent another wave.

They are telling you the script and while many British people believed 'Covid' restrictions would end in the summer of 2021 the government was preparing for them to be ongoing. Authorities were awarding contracts for 'Covid marshals' to police the restrictions with contracts starting in July, 2021, and going through to January 31st, 2022, and the government was advertising for 'Media Buying Services' to secure media propaganda slots worth a potential £320 million for 'Covid-19 campaigns' with a contract not ending until March, 2022. The recipient – via a list of other front companies – was reported to be American media marketing giant Omnicom Group

Inc. While money is no object for 'Covid' the UK waiting list for all other treatment – including life-threatening conditions – passed 4.5 million. Meantime the Cult is seeking to control all official 'inquiries' to block revelations about what has really been happening and why. It must not be allowed to – we need Nuremberg jury trials in every country. The cover-up doesn't get more obvious than appointing ultra-Zionist professor Philip Zelikow to oversee two dozen US virologists, public health officials, clinicians, former government officials and four American 'charitable foundations' to 'learn the lessons' of the 'Covid' debacle. The personnel will be those that created and perpetuated the 'Covid' lies while Zelikow is the former executive director of the 9/11 Commission who ensured that the truth about those attacks never came out and produced a report that must be among the most mendacious and manipulative documents ever written – see *The Trigger* for the detailed exposure of the almost unimaginable 9/11 story in which Sabbatians can be found at every level.

Passive no more

People are increasingly challenging the authorities with amazing numbers of people taking to the streets in London well beyond the ability of the Face-Nappies to stop them. Instead the Nappies choose situations away from the mass crowds to target, intimidate, and seek to promote the impression of 'violent protestors'. One such incident happened in London's Hyde Park. Hundreds of thousands walking through the streets in protest against 'Covid' fascism were ignored by the Cult-owned BBC and most of the rest of the mainstream media, but they delighted in reporting how police were injured in 'clashes with protestors'. The truth was that a group of people gathered in Hyde Park at the end of one march when most had gone home and they were peacefully having a good time with music and chat. Face-Nappies who couldn't deal with the full-march crowd then waded in with their batons and got more than they bargained for. Instead of just standing for this criminal brutality the crowd used their numerical superiority to push the Face-Nappies out of the

park. Eventually the Nappies turned and ran. Unfortunately two or three idiots in the crowd threw drink cans striking two officers which gave the media and the government the image they wanted to discredit the 99.9999 percent who were peaceful. The idiots walked straight into the trap and we must always be aware of potential agent provocateurs used by the authorities to discredit their targets.

This response from the crowd – the can people apart – must be a turning point when the public no longer stand by while the innocent are arrested and brutally attacked by the Face-Nappies. That doesn't mean to be violent, that's the last thing we need. We'll leave the violence to the Face-Nappies and government. But it does mean that when the Face-Nappies use violence against peaceful people the numerical superiority is employed to stop them and make citizen's arrests or Common Law arrests for a breach of the peace. The time for being passive in the face of fascism is over.

We are the many, they are the few, and we need to make that count before there is no freedom left and our children and grandchildren face an ongoing fascist nightmare.

COME ON PEOPLE – IT'S TIME.

One final thought …

The power of love
A force from above
Cleaning my soul
Flame on burn desire
Love with tongues of fire
Purge the soul
Make love your goal

I'll protect you from the hooded claw
Keep the vampires from your door
When the chips are down I'll be around
With my undying, death-defying
Love for you

Envy will hurt itself
Let yourself be beautiful
Sparkling love, flowers
And pearls and pretty girls
Love is like an energy
Rushin' rushin' inside of me

This time we go sublime
Lovers entwine, divine, divine,
Love is danger, love is pleasure
Love is pure – the only treasure

I'm so in love with you
Purge the soul
Make love your goal

The power of love
A force from above
Cleaning my soul
The power of love
A force from above
A sky-scraping dove

Flame on burn desire
Love with tongues of fire
Purge the soul
Make love your goal

Frankie Goes To Hollywood

APPENDIX

Cowan-Kaufman-Morell Statement on Virus Isolation (SOVI)

Isolation: The action of isolating; the fact or condition of being isolated or standing alone; separation from other things or persons; solitariness
Oxford English Dictionary

The controversy over whether the SARS-CoV-2 virus has ever been isolated or purified continues. However, using the above definition, common sense, the laws of logic and the dictates of science, any unbiased person must come to the conclusion that the SARS-CoV-2 virus has never been isolated or purified. As a result, no confirmation of the virus' existence can be found. The logical, common sense, and scientific consequences of this fact are:

- the structure and composition of something not shown to exist can't be known, including the presence, structure, and function of any hypothetical spike or other proteins;
- the genetic sequence of something that has never been found can't be known;
- "variants" of something that hasn't been shown to exist can't be known;
- it's impossible to demonstrate that SARS-CoV-2 causes a disease called Covid-19.

In as concise terms as possible, here's the proper way to isolate, characterize and demonstrate a new virus. First, one takes samples (blood, sputum, secretions) from many people (e.g. 500) with symptoms which are unique and specific enough to characterize an illness. Without mixing these samples with ANY tissue or products that also contain genetic material, the virologist macerates, filters and ultracentrifuges i.e. *purifies* the specimen. This common virology technique, done for decades to isolate bacteriophages[1] and so-called giant viruses in every virology lab, then allows the virologist to demonstrate with electron microscopy thousands of identically sized and shaped particles. These particles are the isolated and purified virus.

These identical particles are then checked for uniformity by physical and/or microscopic techniques. Once the purity is determined, the particles may be further characterized. This would include examining the structure, morphology, and chemical composition of the particles. Next, their genetic makeup is characterized by extracting the genetic material directly from the purified particles and using genetic-sequencing techniques, such as Sanger sequencing, that have also been around for decades. Then one does an analysis to confirm that these uniform particles are exogenous (outside) in origin as a virus is conceptualized to be, and not the normal breakdown products of dead and dying tissues.[2] (As of May 2020, we know that virologists have no way to determine whether the particles they're seeing are viruses or just normal breakdown products of dead and dying tissues.)[3]

---

1 Isolation, characterization and analysis of bacteriophages from the haloalkaline lake Elmenteita, KenyaJuliah Khayeli Akhwale et al, PLOS One, Published: April 25, 2019. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0215734 – accessed 2/15/21

2 "Extracellular Vesicles Derived From Apoptotic Cells: An Essential Link Between Death and Regeneration," Maojiao Li1 et al, Frontiers in Cell and Developmental Biology, 2020 October 2. https://www.frontiersin.org/articles/10.3389/fcell.2020.573511/full – accessed 2/15/21

3 "The Role of Extraellular Vesicles as Allies of HIV, HCV and SARS Viruses," Flavia Giannessi, et al, Viruses, 2020 May

If we have come this far then we have fully isolated, characterized, and genetically sequenced an exogenous virus particle. However, we still have to show it is causally related to a disease. This is carried out by exposing a group of healthy subjects (animals are usually used) to this isolated, purified virus in the manner in which the disease is thought to be transmitted. If the animals get sick with the same disease, as confirmed by clinical and autopsy findings, one has now shown that the virus actually causes a disease. This demonstrates infectivity and transmission of an infectious agent.

None of these steps has even been attempted with the SARS-CoV-2 virus, nor have all these steps been successfully performed for any so-called pathogenic virus. Our research indicates that a single study showing these steps does not exist in the medical literature.

Instead, since 1954, virologists have taken unpurified samples from a relatively few people, often less than ten, with a similar disease. They then minimally process this sample and inoculate this unpurified sample onto tissue culture containing usually four to six other types of material – all of which contain identical genetic material as to what is called a "virus." The tissue culture is starved and poisoned and naturally disintegrates into many types of particles, some of which contain genetic material. Against all common sense, logic, use of the English language and scientific integrity, this process is called "virus isolation." This brew containing fragments of genetic material from many sources is then subjected to genetic analysis, which then creates in a computer-simulation process the alleged sequence of the alleged virus, a so called in silico genome. At no time is an actual virus confirmed by electron microscopy. At no time is a genome extracted and sequenced from an actual virus. This is scientific fraud.

The observation that the unpurified specimen — inoculated onto tissue culture along with toxic antibiotics, bovine fetal tissue, amniotic fluid and other tissues — destroys the kidney tissue onto which it is inoculated is given as evidence of the virus' existence and pathogenicity. This is scientific fraud.

From now on, when anyone gives you a paper that suggests the SARS-CoV-2 virus has been isolated, please check the methods sections. If the researchers used Vero cells or any other culture method, you know that their process was not isolation. You will hear the following excuses for why actual isolation isn't done:

1. There were not enough virus particles found in samples from patients to analyze.

2. Viruses are intracellular parasites; they can't be found outside the cell in this manner.

If No. 1 is correct, and we can't find the virus in the sputum of sick people, then on what evidence do we think the virus is dangerous or even lethal? If No. 2 is correct, then how is the virus spread from person to person? We are told it emerges from the cell to infect others. Then why isn't it possible to find it?

Finally, questioning these virology techniques and conclusions is not some distraction or divisive issue. Shining the light on this truth is essential to stop this terrible fraud that humanity is confronting. For, as we now know, if the virus has never been isolated, sequenced or shown to cause illness, if the virus is imaginary, then why are we wearing masks, social distancing and putting the whole world into prison?

Finally, if pathogenic viruses don't exist, then what is going into those injectable devices erroneously called "vaccines," and what is their purpose? This scientific question is the most urgent and relevant one of our time.

We are correct. The SARS-CoV2 virus does not exist.

Sally Fallon Morell, MA
Dr. Thomas Cowan, MD
Dr. Andrew Kaufman, MD

Bibliography

Alinsky, Saul: *Rules for Radicals* (Vintage, 1989)

Antelman, Rabbi Marvin: *To Eliminate the Opiate* (Zahavia, 1974)

Bastardi, Joe: *The Climate Chronicles* (Relentless Thunder Press, 2018)

Cowan, Tom: *Human Heart, Cosmic Heart* (Chelsea Green Publishing, 2016)

Cowan, Tom, and Fallon Morell, Sally: *The Contagion Myth* (Skyhorse Publishing, 2020)

Forbes, Jack D: *Columbus And Other Cannibals – The Wetiko Disease of Exploitation, Imperialism, and Terrorism* (Seven Stories Press, 2008 – originally published in 1979)

Gates, Bill: *How to Avoid a Climate Disaster: The Solutions We Have and the Breakthroughs We Need* (Allen Lane, 2021)

Huxley, Aldous: *Brave New World* (Chatto & Windus, 1932)

Köhnlein, Dr Claus, and Engelbrecht, Torsten: *Virus Mania* (emu-Vertag, Lahnstein, 2020)

Lanza, Robert, and Berman, Bob: *Biocentrism* (BenBella Books, 2010)

Lash, John Lamb: *Not In His Image* (Chelsea Green Publishing, 2006)

Lester, Dawn, and Parker, David: *What Really Makes You Ill – Why everything you thought you knew about disease is wrong* (Independently Published, 2019)

Levy, Paul: *Dispelling Wetiko, Breaking the Spell of Evil* (North Atlantic Books, 2013)

Marx, Karl: *A World Without Jews* (Philosophical Library, first edition, 1959)

Mullis, Kary: *Dancing Naked in the Mine Field* (Bloomsbury, 1999)

O'Brien, Cathy: *Trance-Formation of America* (Reality Marketing, 1995)

Scholem, Gershon: *The Messianic Idea in Judaism* (Schocken Books, 1994)

Schwab, Klaus, and Davis, Nicholas: *Shaping the Future of the Fourth Industrial Revolution: A guide to building a better world* (Penguin Books, 2018)

Schwab, Klaus: *The Great Reset* (Agentur Schweiz, 2020)

Sunstein, Cass and Thaler, Richard: *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Penguin, 2009)

Swan, Shanna: *Count Down: How Our Modern World Is Threatening Sperm Counts, Altering Male and Female Reproductive Development and Imperiling the Future of the Human Race* (Scribner, 2021)

Tegmark, Max: *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality* (Penguin, 2015)

Velikovsky, Immanuel: *Worlds in Collision* (Paradigma, 2009)

Wilton, Robert: *The Last Days of the Romanovs* (Blurb, 2018, first published 1920)

Index

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

Y

Z

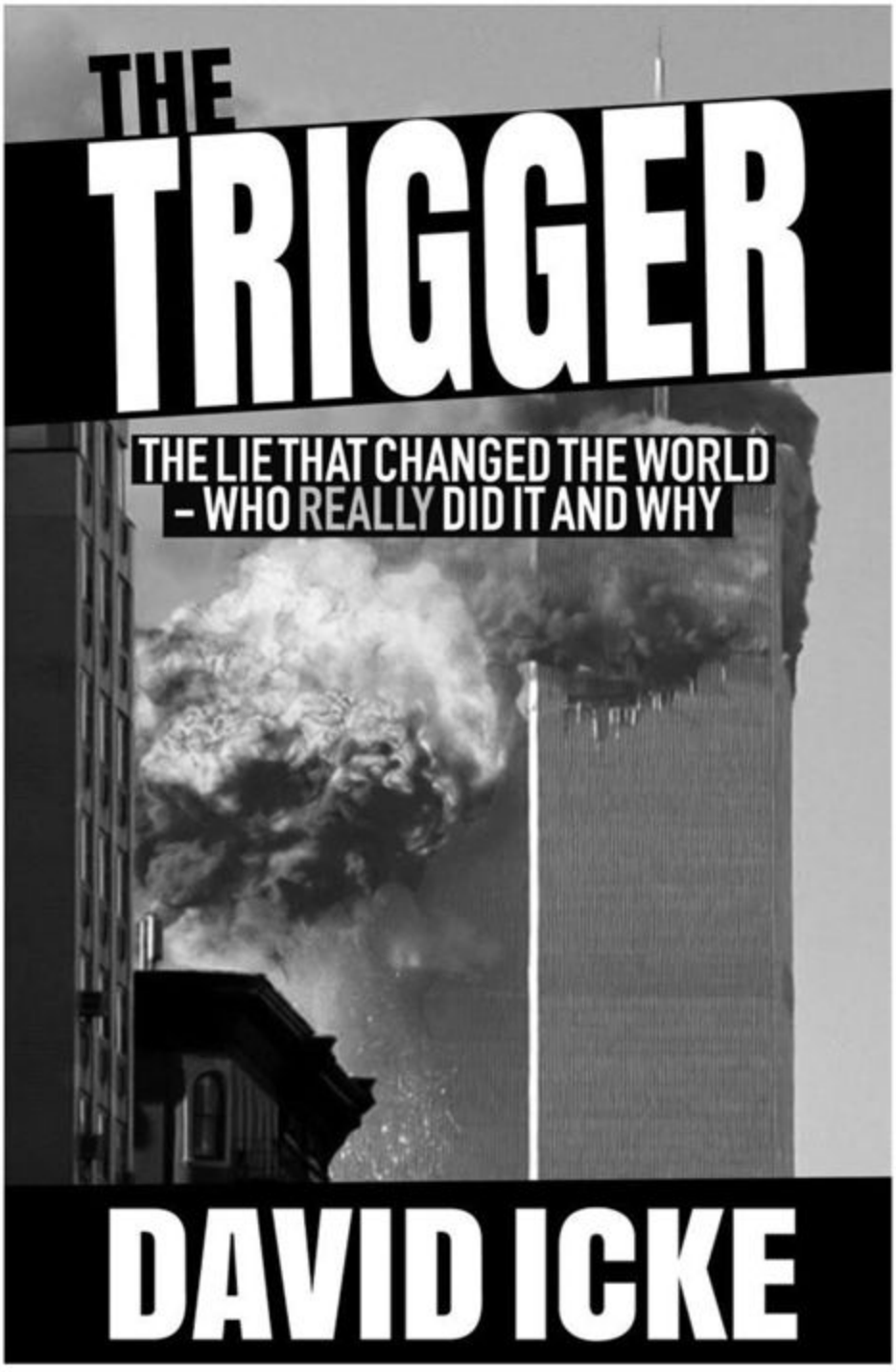
ICKONIC
THE ALTERNATIVE

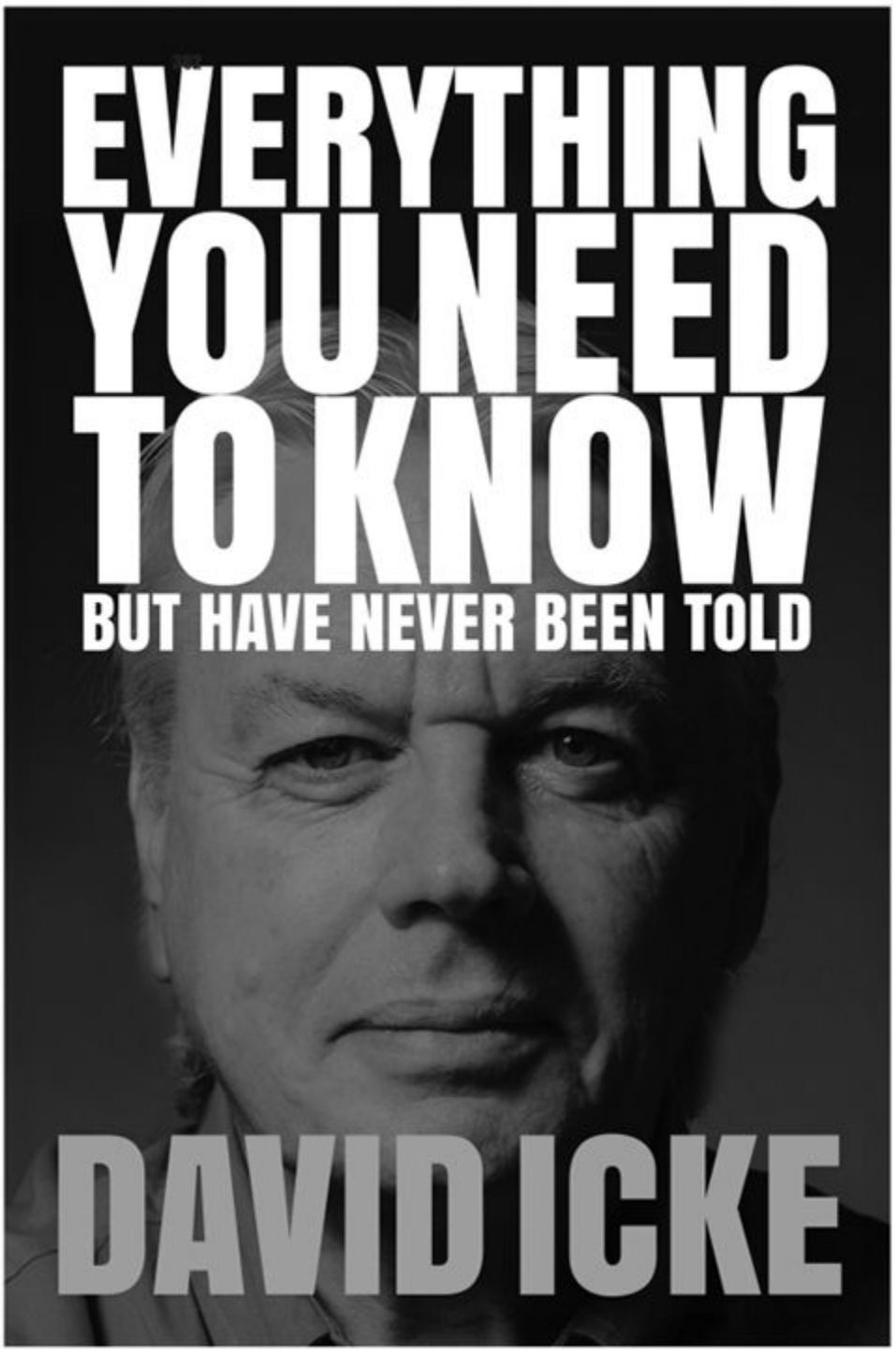
Ickonic is something that has been a dream of mine for the last 5 years, growing up around alternative information I have always had a natural interest in what is going on in the World and what could I do to make it better.
Across the range of subjects and positions of influence occupied mainly by people who don't strive to make things better it's the Media that I have always found the most frustrating and fascinating. Mainly because if the Media did their Jobs properly then so much of the negative things happening in the World simply would not be able to happen, because they would be exposed within a heartbeat.
Free Press and the Opportunities that the internet could have given would mean that the Media are able to expose things like never before and hold people to account for their actions. As we all know there are 'Untouchables' that walk among us, people the Media simply won't touch, expose or investigate and that leads to the dark underworlds that infest the establishment the World over.
Well I say enough, it's time for something different, a different kind of Media, where no one is off limits from exposing and investigating. All we're interested in at Ickonic is the truth of what is really going on in the World on whichever subject we're covering.
We hope you enjoy what we have created and take something away from the platform, we aim to deliver information that's informative and most importantly self-empowering, you're not a little person, you're part of something much bigger than that and its time we as a collective race began to understand that and look to the future as ours to take.

It's time...

Jaymie Icke - Founder Ickonic Alternative Media.

SIGN UP NOW AT ICKONIC.COM


THE
TRIGGER
THE LIE THAT CHANGED THE WORLD
– WHO REALLY DID IT AND WHY
DAVID ICKE


EVERYTHING
YOU NEED
TO KNOW
BUT HAVE NEVER BEEN TOLD
DAVID ICKE

DAVIDICKE.COM

DAVID ICKE STORE

LATEST NEWS ARTICLES

DAVID ICKE VIDEOS

WEEKLY DOT-CONNECTOR PODCASTS

LIVE EVENTS

WWW.DAVIDICKE.COM

Before you go ...

For more detail, background and evidence about the subjects in *Perceptions of a Renegade Mind* – and so much more – see my others books including *And The Truth Shall Set You Free; The Biggest Secret; Children of the Matrix; The David Icke Guide to the Global Conspiracy; Tales from the Time Loop; The Perception Deception; Remember Who You Are; Human Race Get Off Your Knees; Phantom Self; Everything You Need To Know But Have Never Been Told, The Trigger and The Answer*.

You can subscribe to the fantastic new Ickonic media platform where there are many hundreds of hours of cutting-edge information in videos, documentaries and series across a whole range of subjects which are added to every week. This includes my 90 minute breakdown of the week's news every Friday to explain *why* events are happening and to what end.